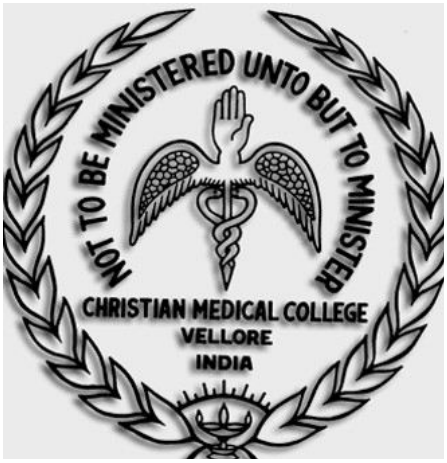
A STUDY ON THE VALIDATION OF PAEDIATRIC INDEX OF MORTALITY 2 (PIM2) SCORE IN THE PAEDIATRIC INTENSIVE CARE UNIT OF CMC, VELLORE



A dissertation submitted to the TN. Dr. MGR medical University, Tamil Nadu, in partial fulfillment of the rules and regulations for the M.D. branch VII (PAEDIATRICS) Examination to be held in March 2008

CERTIFICATE

This is to certify that this dissertation entitled

"A STUDY ON THE VALIDATION OF PAEDIATRIC INDEX OF MORTALITY 2 (PIM2) SCORE IN THE PAEDIATRIC INTENSIVE CARE UNIT OF CMC, VELLORE"

is a bonafide work done by

Dr.Muthulakshmi.N.

in partial fulfillment of the rules and regulations for the M.D. Branch VII (PAEDIATRICS) examination of the Tamilnadu, Dr. M.G.R. Medical University, Chennai, to be held in March 2008.

Dr. KALA EBENEZER M.D., D.C.H
Professor,
Division of Paediatric Intensive Care,
Department of Child Health,
Christian Medical College,
Vellore. Tamil Nadu

Dr. ATANU KUMAR JANA M.D., D.C.H
Professor and Head,
Department of Child Health,
Christian medical College,
Vellore, Tamil Nadu

PLACE:

DATED:

Declaration

I hereby declare that the investigations that form the subject matter of my thesis was carried out by me under the guidance of Dr. Kala Ebenezer M.D., D.C.H, Professor, Division of Paediatric Intensive Care, Department of Child Health, Christian Medical College,Vellore. This has not been submitted in any other university in part or in full.

Dr.Muthulakshmi.N
Department of Child Health
Christian Medical College
Vellore.

Place: Vellore.

ACKNOWLEDGEMENTS

I wish to express my deep gratitude to Dr. Kala Ebenezer, M.D., D.C.H., Professor and Head of the Paediatric Intensive Care Unit, and Professor, Department of Child Health, Christian Medical College, Vellore, for her valuable guidance and constant encouragement throughout the course of the study.

I am thankful to Dr. Atanu Kumar Jana, M.D., D.C.H., Professor and head of the Department of Child Health, for his interest and help in the successful completion of the study.

I am extremely thankful to all patients, my colleagues and staff of the Paediatric Intensive Care Unit for the cooperation they extended during the course of the study.

I wish to express my deep appreciation to Mr. Solomon, Department of Biostatistics and Dr. Amit Kumar Datta, M.D, D.M, Department of Gastro Enterology, for their valuable help in the statistical analysis of the study.

CONTENTS

AIMS AND OBJECTIVES

- To validate the Paediatric Index of Mortality 2 (PIM 2) score in the paediatric intensive care unit of CMC, Vellore
- To compare the performance of our PICU with the other intensive care units in the world.

Introduction

Paediatric intensive care is a rapidly developing super specialty especially in the later half of the 20th century. While these developments happened much later in developing countries, over the past 10 years there has been a tremendous growth in Paediatric intensive care in India.

Paediatric Intensive Care Units (PICU) constantly aim at promoting care to critically ill children which often involves huge amounts of technology that comes at exorbitant cost, often unreachable to many families. Even though the underlying condition is potentially treatable, the outcome is often uncertain.

The question of how best to utilize critical care resources is one that challenges intensive care unit (ICU) directors on a daily basis. Shortages in space, human and economic resources especially in developing countries limit the ability to provide complete care. As costs of health care in ICU have gone up dramatically, correctly identifying those children who are salvageable from moribund group of children becomes necessary for implementation of effective and rational medical therapy.

The Paediatric Intensive care unit (PICU) at Christian Medical College and Hospital, Vellore is a multispecialty unit that provides services to a heterogeneous group of sick children from various paediatric specialties including General Paediatrics, Paediatric oncology, Paediatric Nephrology, Paediatric Endocrinology, Developmental Paediatrics, Paediatric surgery, Hematology, Urology, Plastic surgery and Cardiology. The annual admission is approximately 1200 children per year from various parts of India mainly comprising population from in and around Vellore including adjacent parts of Andhra Pradesh. Since this is a major tertiary center, most admissions are referred from various nursing homes as well as medical college hospitals for further management in view of disease severity. About one-fourth of PICU admissions are post- operative patients from surgical specialties. Thus our patient population represents a good amount of case mix and disease severity requiring intensive care.

Infectious diseases, respiratory and sepsis syndromes comprise the majority of admissions followed by neurological illness. Majority of the patient population are infants. The overall mortality is about 20-25% which is high compared to figures from intensive care units abroad. This finding is probably related to the patient profile as well as a greater load of sicker patients being managed with scarce resources.

This has compelled us to do an outcome analysis by predictive scoring system so that accurate outcome data is available for guidance of prognostication and counseling the parents.

Literature review had shown that there are two important scores that are available for use in Paediatric intensive care units. The Paediatric Risk of Mortality score (PRISM) published in 1988 by Pollack et al. is still the most widely known and used prognostic score for evaluation of disease severity in children. In 1997 Shann et al developed and validated the Paediatric Index of Mortality score (PIM SCORE) which was simpler to collect as it had less variables than PRISM and performed well as compared to PRISM. The same authors published a revised score in 2003- PIM-2 which was better calibrated, safer and better adjusted for diagnostic groups than its original version.

The PIM Score which was originally developed in Australian PICUs, has been tested and found to have good discriminatory capacity in intensive care units in the U.K and by the same group of authors. Other studies from Hong Kong, Argentina and Italy have also shown comparable results. Limited Indian data is available in the literature. The only study from India compared the performance of Pediatric Risk of Mortality (PRISM), Pediatric Index of mortality (PIM) and PIM2 in 246 children admitted to a 6 bedded PICU in New Delhi. The area under the receiver operating characteristic curve for all models was >0.8. However all scores under predicted mortality.

We expected the patients in our ICU to be quite different from that of Australian intensive care units in terms of the underlying disease, the severity of their illness, the access to health care, the amount of intensive care available and the health seeking pattern which is influenced by the cultural background from where they are from.

Moreover, nutritional status of the children will be very different and will be a major compounding factor that might influence the outcome.

We were interested in determining if the PIM score is applicable for our intensive care unit with its patients as described. PIM data is easy to collect and is available on the internet free of cost. It is done at time of admission and is independent of the intensive care delivered to the patients. Hence this study of validating the PIM-2 score in the PICU of Christian Medical College & Hospital was undertaken.

Literature Review

The provision of intensive care to infants, children, and adults increased steadily during the latter half of the 20th century, with particular, rapid expansion during the 1970s and 1980s. In paediatric intensive care, the growth in activity has followed somewhat later.

The main goals for intensive care can be simplified to save the lives of the salvageable patients with reversible medical conditions and offer the dying patient a peaceful and dignified death. Correctly identifying these two groups of patients is necessary for implementation of effective and rational medical therapy. The escalation in health care costs over the past 25 years has paralleled the development of specialized intensive care units and the recognition of critical care medicine as a separate medical specialty. ICU beds make up about 5% -10% of the total hospital beds. In developed countries such as the United States of America the total health care costs equal approximately 11-12% of GNP with the total critical care costs representing about 1% of the GNP[1]. Similar data in the developing world is forthcoming.

Thus outcome analysis by predictive scoring systems has become a growing field of interest in clinical research conducted in critical care medicine. Two important reasons for this are the rising cost of health care and the need for accurate outcome data so that patients, their families, and physicians can make informed health care decisions.

Utilization of the intensive care unit

The question of how best to utilize critical care resources is one that challenges intensive care unit (ICU) directors on a daily basis. Shortages in space, human and economic resources limit the ability to provide complete care. As a result, some form of selection is necessary to allocate these scarce resources among the individuals competing for them – a process termed rationing [2].

Rationing of ICU facilities

Rationing of all types can occur in ICU setting – rationing of equipment, staffing, medications especially as ICU is the site of application of most advanced technologies. Admission to ICU often serves as a signal to proceed with a continuous series of interventions – a process that can be difficult to limit or stop.

Decisions about which patients should be admitted and which patients should be rejected may consume relatively large amounts of time; in fact, decisions about how to manage a patient after admission to the ICU often are less labored than the discussions about whether or not to admit the patient to ICU.

The imbalance between the supply and the demand for ICU beds and the resulting rationing that must occur justify the need to develop formal guidelines for ICU admission and allocation of this scarce resource[2]. This is particularly the case as patient selection for ICU admission often must be accomplished expeditiously, on short notice, at inopportune times and by junior members of the ICU staff. Unless specific policy guidelines are developed, implemented and adhered to such decision making becomes subject to the variabilities of personal biases.

Establishing ICU guidelines is not as simple as it might seem and the process itself reveals ethically troublesome issues. The information available for sound decision making is expanding. Protocols to evaluate the outcomes of patients suffering specific medical problems have led to the development of scoring systems for objectively quantitating severity of illness and predicting the outcomes based on specific underlying diseases and physiologic conditions[3,4,5]. Although this information has improved our knowledge base, the incorporation of this knowledge into decision-making process of ICU management remains complex.

Scoring systems versus randomized controlled trials

Although rigorous experiments or large randomized controlled trials are the gold standard for evaluating existing or new interventions, these are not always possible in intensive care. For example, it is unethical to randomly allocate severely ill patients to receive intensive care or general ward care**.** The alternative is to use observational methods that study the outcome of care patients receive as part of their natural treatment[6].

Evaluation of Outcome

Intensive care has developed over the past 30 years. Without enough rigorous scientific evidence on clinical evidence, doctors delivering intensive care often have to decide which patients will benefit most. Scoring systems have been developed in response to an increasing emphasis on the evaluation and monitoring of health services. These systems enable comparative audit and evaluative research of intensive care[6].

The evaluation of any outcome requires performing two specific processes. The First, the outcome of interest should be clearly defined so that it can be measured accurately and second, methods or techniques should be available to predict the outcome[1]. In most outcome investigations involving critically ill patients, the outcome measured has been patient mortality. Other outcomes representing important societal issues have not received the same attention as mortality in critical care but include the effectiveness of resource utilization and cost benefit analysis of therapies and diagnostic modalities used in critical care, the resultant attitudes of patients and their families in regard to critical care, the long term quality of life after ICU care and the systematic evaluation of the process of critical care as it is being administered[1].

Risk Adjusted Mortality

Risk adjusted mortality remains the commonest benchmark for neonatal, paediatric, and adult ICU performance[7]. Mortality risk scoring systems are integral to the provision of modern intensive care, providing a measure of performance both between and within individual intensive care units over time. However, before inferences can be

drawn about outcomes of treatment in such studies the characteristics of the patients admitted to intensive care have to be taken into account. This process is known as adjusting for case mix. Scoring systems are aimed at quantifying case mix and using the resulting score to estimate outcome.

The death rate of patients admitted to intensive care units is much higher than that of other hospital patients. Given the relatively high mortality among intensive care patients, death is a sensitive, appropriate, and meaningful measure of outcome. However, death can result from many factors other than ineffective care. Outcome depends not only on the input (equipment, staff) and the processes of care (type, skill, and timing of care) but also on the case mix of the patients. The patient population of an intensive care unit in a large tertiary care centre may be very different from that of a unit based in a district general hospital. Patients are admitted to intensive care for a wide range of clinical indications; both the nature of the current crisis and any underlying disease must be considered. Intensive care units admitting greater proportions of high risk patients would be expected to have a higher mortality.

A valid scoring system must predict mortality accurately while adjusting for case mix and disease severity, but also requires data capture that is feasible in clinical practice, and should be updated to reflect changes and advances in ICU care[7].

Measurement of Outcome

The essential characteristics of an ideal predictor of outcome would include

- The prediction rule should have a well defined biological outcome that is easily measured and clinically relevant.
- The rule should be derived from a broad, large database so that it would be applicable to a large number of patients.
- The error rate of any prediction rule should be well known and its significance well understood.

The likelihood of mortality assigned to a patient by any predictive process is termed a probability estimate, of which there are two types - Subjective probability

founded entirely upon the knowledge of physicians and objective probability, drawn from databases of compiled clinical information.

Criteria for selecting a scoring system

- Proposed use
- Validity of score
- Reliability of score
- Discrimination of scoring system
- Calibration of scoring system

Evolution of scoring systems in intensive care

Paediatric scoring systems can be broadly classified as in table.1. The earlier scoring systems[6] were developed for trauma patients and were either

specific anatomical methods	Abbreviated injury score, 1969
	Burns score, 1971
	Injury severity score, 1974
or	
specific physiological methods	Trauma index, 1971
	Glasgow coma scale, 1974
	Trauma score, 1981
	Sepsis score, 1983

A brief description of important scoring systems which are in vogue follows. Description of scoring systems pertaining to the study will follow.

The Abbreviated Injury Score

The Abbreviated Injury Scale[8] (AIS) is an anatomical scoring system first introduced in 1969. Since this time it has been revised and updated against survival so that it now provides a reasonably accurate ranking of the severity of injury. The latest incarnation of the AIS score is the 1990 revision (Table 2). The AIS is monitored by a scaling committee of the Association for the Advancement of Automotive Medicine.

Injuries are ranked on a scale of 1 to 6, with 1 being minor, 5 severe and 6 an unsurvivable injury. This represents the 'threat to life' associated with an injury and is not meant to represent a comprehensive measure of severity. Organ Injury Scales of the American Association for the Surgery of Trauma are mapped to the AIS score for calculation of the Injury severity score.

Injury severity score (1974)

The Injury Severity Score[9] (ISS) (Table 3) is an anatomical scoring system that provides an overall score for patients with multiple injuries. Each injury is assigned an Abbreviated Injury Score (AIS) score, allocated to one of six body regions (Head, Face, Chest, Abdomen, Extremities (including Pelvis), External). Only the highest AIS score in each body region is used. The 3 most severely injured body regions have their score squared and added together to produce the ISS score.

The ISS takes values from 0 to 75. An example of the ISS calculation is shown in table 3. If an injury is assigned an AIS of 6 (unsurvivable injury), the ISS score is automatically assigned to 75. The ISS correlates linearly with mortality, morbidity, hospital stay and other measures of severity.

It's weaknesses are that any error in AIS scoring increases the ISS error, many different injury patterns can yield the same ISS score and injuries to different body regions are not weighted. Since a full description of patient injuries is not known prior to full investigation & operation, the ISS (along with other anatomical scoring systems) is not useful as a triage tool.

Numerous other predictive systems have been developed to forecast patient outcomes in many clinical setting. Examples are the APACHE (Acute Physiology And Chronic Health Evaluation), the MPM (mortality prediction model), and the SAPS (Simplified Acute Physiological Score) aimed at the general intensive care population. Apart from these, models have also been developed to focus on the paediatric age group namely the PRISM, PRISM II, PRISM III, PELOD, PIM and the PIM2 which are discussed further on. For the neonates, scores such as the neonatal therapeutic scoring system (N-TISS)[10], Clinical Risk Index For Babies (CRIB)[11], Simplified Newborn Illness Severity and Mortality Risk Scores (SNAP)[13] the has been in use in neonatal centers.

Measuring severity by treatment

Therapeutic Intervention Scoring System (TISS)

The therapeutic intervention scoring system[3] (TISS) published in 1974 was developed to quantify severity of illness among intensive care patients based on the type and amount of treatment received. The underlying philosophy was that the sicker the patient, the greater the number and complexity of treatments given. By quantifying this, a proxy measure of the severity of illness for a patient could be obtained. The system scored 76 common therapeutic activities and was last updated in 1983. A simplified version based on 28 therapeutic activities (TISS 28)[14] has been published and a version for patients in high dependency units has been proposed. Another approach is to assess the severity of organ dysfunction based on the type and amount of treatment received. These organ failure scoring systems are used to give a probability of hospital death which takes into account the severity of dysfunction in each organ system and the effect on prognosis of dysfunction in several organ systems.

Objective probability estimates using continuous variables

Acute Physiology Score and Chronic Health Evaluation (APACHE)

In the mid1970s William Knaus developed the APACHE[15]. This score uses continuous variables such as body temperature, serum sodium and pulse rate which are likely to be deranged in the setting of critical illness and therefore are more likely to portray an accurate image of the severity of physiologic derangement than non continuous variables.

However the large number of variables required for the APACHE I became a source of systemic inaccuracy. The APACHE I[15] presented a practical problem because there were 34 variables to be measured and certain variables such as serum osmolality and serum lactate were omitted more frequently than the others. In this model some variables deemed important were often not measured, with the problem being confounded by the model considering missing data as normal. Following a multivariate analysis involving each of the apache score variables and mortality it was found that the variables

considered critical did not improve the overall explanatory power of the system. A major criticism of the original APACHE system was that the variables were chosen by a group of physicians. This panel of experts approach to selection of variables introduces the potential for bias.

Apache II

The APACHE II[4,15] (1985)(Table 4) further reduced the number of variables to 12. When judged against the ideal predictive instrument the APACHE II has several strengths. It had a well defined outcome (Hospital Death); it was derived from a large data base, and the source of bias present in its prototype was understood and corrected. It takes into account the age, pre morbid conditions, and the principal reason for ICU admission. It is accurate in predicting overall mortality in critically ill patients and has been used widely in clinical trials.

However the APACHE is not free of disadvantages. The scoring occurs within 24 hours after admission and there is no adjustment for the clinical course. The important sources of error and bias in the APACHE II include a selection bias where inaccuracies were revealed while predicting mortality in patients with trauma (due to inadequate weight for Glasgow coma scale), cardiogenic pulmonary edema, low albumin states, hematological malignancy and hepatic failure. The acute physiologic score does not take into account previous treatment and hence may underestimate the mortality of patients referred from other hospitals and intensive care units. Another source of error would be the reliance on a single diagnosis in the regression equation in calculating the risk of mortality. Patients in the ICU often suffer from simultaneous conditions involving more than one organ system. Several errors with relation to data collection also plague the APACHE II, with the worst values over the first 24 hours of ICU stay being required. Considerable inter-observer variability with relation to detail is also a contributing factor.

Apache III

The APACHE III[15] draws from a larger and more diverse patient population than the APACHE II. The patient's location prior to ICU admission has been incorporated into the scoring to account for lead time bias. A subtle but important difference between the two is that the APACHE III does not allow accurate mortality prediction within the disease specific group of patients. Rather this score can be used to compare disease

severity between patients suffering form the same disorder. Even so the presence of a selection bias is unavoidable as ICU admission is an arbitrary clinical decision.

The majority of predictive power derives from weighted continuous physiological variables used to calculate a physiology score. When combined with additional points for age and significant co morbid conditions a raw score is obtained. The raw score can range from 0 to 299 and there is a statistically significant increase in predicted mortality for a given diagnosis with each five point increase. New variables such as albumin, bilirubin, BUN, blood glucose levels and urine output values were added.
Recently the APACHE IV[16] has been devised to help quantify severity of illness better.

***Simplified Acute Physiology Score* (*SAPS*)**

The Simplified Acute Physiology Score[15] was developed in response to the complexity and the time consuming nature of the APACHE I. In this model the 34 variables that formed apart of the APACHE score were subjected to a discriminate analysis and multiple linear regression test, from which 13 variables were identified as having the most discriminate power and were the most frequent ones measured covering all major systems. Elements of SAPS have been incorporated into subsequent models such as the SAPS II (1993).

The Mortality Prediction Model (MPM)

Separate predictive equations were developed for estimates of mortality risk at the time of ICU admission, at 24 hours and after 48 hours of admission[15]. Eleven clinical variables are entered. In addition MPM mortality predictions are not dependant on admission diagnosis thus eliminating a source of error. All the information required to calculate the MPM score is available at the time of ICU admission. Therefore the MPM can be used immediately after admission as against the APACHE which has to be done 24 hours. However most of the variables are non continuous in nature and misclassification can lead to a large error. The other limitation of the MPM was that its data base was initially drawn up from a single hospital raising the possibility of a selection bias.

Glasgow Coma Scale is a score which is still in general use in intensive care. The scale avoids having to describe a patient's level of neurological function in words and the assumption that colleagues understand the same meaning from those words. A similar scale used in children is shown in table 5.

Scores Specific for Paediatric Intensive Care

They enable us to investigate the best ways of organizing paediatric intensive care (by comparing different units), to monitor the effects of changes in practice (by observing trends within units over time), to assess the relationship between severity of illness and length-of-stay or cost, and to monitor the effects of rationing intensive care[17].

PRISM, PIM, and PELOD scores are **composite scores**[18] (aggregate scales) that are made up of a group of variables. It is the death rate that was used as the outcome measure to estimate the validity of all these scores. Many types of variables can be used in constructing such scores, including clinical data like heart rate, physiologic data like cardiac index, laboratory data like creatinine or PaO2, and other scores like the Glasgow coma score that is integrated into the PRISM score. The number of points of each variable are be proportional to its capacity to predict a given outcome. More points are attributed to a given variable if the predictive value of the organ or system monitored by this variable is more significant. More points are given if the dysfunction is more severe. For example, in the PELOD[12] score, severe dysfunction of the cardiovascular or neurologic system is more heavily weighted (up to 20 points) than the renal system (maximum of 10 points). On the other hand, 20, 10, 1, 0 points can be given in the PELOD score for more or less severe dysfunction of the cardiovascular or neurologic system.

Prognostic Scores versus Outcome Scores

Prognostic scores were developed to better describe the severity of illness at baseline of groups of critically ill patients. These scores consider some co-morbidities and physiologic disturbances at entry into the pediatric intensive care unit (PICU) or at randomization in a clinical trial. They were developed to maximize prediction of the overall risk of mortality among groups of critically ill patients, given the severity of the patients.

On the other hand outcome scores describe the severity of illness during stay in the intensive care unit (ICU). In this instance, organ physiologic disturbances are collected daily from baseline to outcome or discharge from the ICU. The most abnormal values are retained. The total of points can be then computed in risk of mortality. Outcome scores were developed and validated to maximize description of the clinical course of groups of patients. Mean risk of mortality of a population can also be compared with actual mortality to get a standardized mortality ratio.

Relevance of composite scores

Discrimination and calibration are two very important characteristics of a score. Discrimination is the ability of a test to differentiate patients who meet the outcome (for example, death) and those who do not. The discrimination capacity (predictor performance) of a test is best described by its area under the receiver operating characteristics curve. The calibration of a score is the degree of correspondence at different levels of probability between the probability of the outcome (for example, death) as predicted by the score and the observed frequency of the outcome. The statistical question is: are discrepancies between observed and expected mortality statistically significant?

Well validated composite scores can be used to harden soft data. Death is a good example of hard data: it is easy to get a consensus on its diagnosis, and there is no interobserver variability. For example, one can use a qualitative scale to describe the severity of MODS in PICUs: critically ill children can have no organ dysfunction at all or

light, moderate, or severe MODS. The problem with such a 4-grade qualitative scale (no MODS, light, moderate, or severe MODS) is that the interrater variability is large. Such qualitative scale can be considered as soft data because what is meant by words such as light, moderate and severe can be very different from one caregiver to the other. There is indeed strong evidence that qualitative expressions like that are not reliable. This must apply to qualitative scales.

A semi quantitative or ordinal score is clearly better[18]. For example, one can describe the severity of cases of MODS by reporting the number of dysfunctional organs, which can range from zero to six in critically ill children. There are problems here, too, because the risk of death is different from one organ to the other. For example, in a group of critically ill patients, neurologic or cardiovascular dysfunctions are more important and more predictive of death than hepatic dysfunction. A well developed and well-validated quantitative score can take into account the independent weight of each variable that is integrated into it. Greater of fewer points are attributed to each variable included in the score.

Composite scores are relevant if they are used. This is the case when PRISM and PIM scores are frequently used to compare the efficacy of different PICUs, given the expected mortality predicted by these scores in these units (quality assurance and quality assessment), and in clinical trials to compare the severity of illness of patients at randomization.

Subsequently a few important paediatric outcome scoring systems will be reviewed (PELOD, P-MODS, DORA, CRIB, SNAP) followed by a review of the prognostic scoring systems (PTS, MSSS, PRISM, PIM,).

Outcome scores

Paediatric Logistic Organ Dysfunction (PELOD)

Multiple organ dysfunction syndrome is more frequent than death in paediatric intensive care units. Estimation of the severity of this syndrome could be a useful additional outcome measure in clinical trials in such units. The initial study[12] was a prospective, observational, multi centre cohort study in seven multidisciplinary, tertiary-

care paediatric intensive care units of university-affiliated hospitals (two French, three Canadian, and two Swiss) involving 1806 consecutive patients (median age 24 months; inter quartile range 5–90).

PELOD score[12] (1999) includes six organ dysfunctions and12 variables and was recorded daily. For each variable, the most abnormal value each day and during the whole stay were used in calculating the dPELOD (daily PELOD) and PELOD scores, respectively. The outcome was vital status at discharge.

Findings 21% of patients had no organ dysfunction, 26% had one, 25% had two, and 28% had three or more. Case fatality rate was 6·4%. The PELOD score was significantly higher in non-survivors than survivors ($p<0·0001$)[12]. Calibration of PELOD and dPELOD scores was good. Interpretation PELOD and dPELOD scores was found to result in valid outcome measures of the severity of multiple organ dysfunction syndrome in paediatric intensive care units, and their use could reduce the sample size required to complete clinical trials in critically ill children.

Paediatric Multiple Organ Dysfunction Score (P-MODS)

The P MODS[29] (2005) was developed in a pediatric intensive care unit at a tertiary care pediatric teaching hospital. A total of 6,456 pediatric consecutive admissions (mean age 4.62 yrs) admitted to the pediatric intensive care unit were followed up to identify variables that could define organ dysfunction in children and also to develop a Multiple Organ Dysfunction Score (P-MODS) while looking for a correlation of the score with outcome at pediatric intensive care unit discharge.

Patients were randomly separated into two groups: a development set to create the scoring system and a validation set to evaluate score performance and reproducibility. Survivors and non survivors were compared to define variables that were significantly more abnormal in non survivors. Those variables were correlated with pediatric intensive care unit mortality rate. Optimal intervals for each variable were defined on the development set, and their performance was evaluated in the validation set. Descriptors for organ dysfunction were identified in five organ systems: cardiovascular (lactic acid), respiratory (Pa(O(2))/Fi(O(2)) ratio), hepatic (bilirubin), hematological (fibrinogen), and renal (blood urea nitrogen). A grading scale for each variable was set from 0 to 4,

corresponding to mortality rates of <5% and >50%, respectively. P-MODS were calculated by summing the worst score for all variables.

The score correlated strongly and in a graded fashion with pediatric intensive care unit mortality rate. In both sets (development and validation), mortality rate was <5% when the score was 0 and >70% at the highest score. Overall mortality rate was 5.9% (development set) and 5.3% (validation set). The score showed excellent discrimination reflected in areas under the curve: 0.81 (development set) and 0.78 (validation set).

Dynamic Objective Risk Assessment (DORA)

DORA (1991) is a mortality risk predictor based on physiologic data that estimates daily the probability of a patient dying within the next 24 hrs as that probability changes with disease and recovery. The study was initially done in nine pediatric ICUs. Data from 1,401 patients (116 deaths, 5,521 days of care) were used for predictor development, and 1,227 patients (105 deaths, 4,597 days of care) provided data for predictor validation[31].

The predictor was developed by logistic regression analysis using the Paediatric Risk of Mortality (PRISM) scores of all previous days as potential predictor variables. Performance was measured by the area under the receiver operating characteristic curve and by the comparison of the daily predicted vs. observed patient status in five mortality risk groups (less than 0.01, 0.01 to 0.05, 0.05 to 0.15, 0.15 to 0.3, greater than 0.3).

Only the most recent and the admission day Paediatric Risk of Mortality scores (with a weighting ratio of 3:1) contributed significantly (p less than .05) to the prediction. The daily number and distribution of survivors and non survivors in the five mortality risk groups were well predicted in the total sample and each ICU separately. This dynamic predictor improved (p less than .01) ICU outcome prediction over an admission-day predictor[31].

The predictor was previously considered valid for assessing the 24-hr mortality risk in pediatric ICU patients hospitalized in other tertiary care institutions, different from those used for predictor development. The predicted mortality risks allow prospective patient stratification into risk groups. The ability of this predictor to follow risk changes

over time expanded its applicability over static predictors by enabling the charting of patient courses, and permitting ICU efficiency analysis

Clinical Risk Index for Babies (CRIB I and II)

CRIB (Table 6) is a neonatal scoring system. Scores are given for birth weight, gestational age, maximum and minimum fraction of inspired oxygen and maximum base excess during the first 12 h, and presence of congenital malformations[11]. CRIB was developed retrospectively in a cohort of 812 infants of birth weight 1500 g or less or gestational age less than 31 weeks treated in four UK tertiary hospitals between 1988 and 1990. The area under the receiver operating characteristic (ROC) curve for predicting death in this validation cohort (a measure of the predictor's accuracy) was significantly greater for CRIB than for birth weight alone (0·90 [SE 0·05] vs 0·78 [0·03], P=0·03).

Calculation of risk- adjusted mortality by means of CRIB showed that babies were twice as likely to die in the non-tertiary hospitals as in the tertiary hospitals in the UK (odds ratio 2·12 [95% CI 1·39-3·24]). Adjustment for birth weight alone also showed higher odds of death in non-tertiary hospitals (1·45 [1·01-2·11]). CRIB was considered a robust index of initial neonatal risk that is more accurate than birth weight and simple enough for routine use.

The appropriateness of CRIB with contemporary data was subsequently questioned, that it might be no better in prediction of mortality than birth weight or gestation alone. Furthermore, CRIB included fraction of inspired oxygen (FiO_2), which is not a true physiological measure because it is determined by the care team and also includes data up to 12 h after admission, thus potentially introducing early treatment bias.

Therefore a new five-item CRIB II score[32] was developed with data from a UK-wide sample of 1886 infants admitted in 1998–99. CRIB II avoided the potential problems of early treatment bias. CRIB II developed and assessed for infants of 32 weeks' gestation, with further data available up to 1 h after admission to neonatal intensive care, excluding FiO_2.

Simplified Newborn Illness Severity and Mortality Risk Scores
(SNAP, SNAP-II and SNAPPE-II)

Substantial variation in birth weight-adjusted mortality among neonates reflects differences in population illness severity. The Score for Neonatal Acute Physiology[33] (SNAP) was developed in 1992 and validated prospectively on 1643 admissions (114 deaths) in three ICUs. SNAP scored the worst physiologic derangements in each organ system in the first 24 hours. SNAP showed little correlation with birth weight and was highly predictive of neonatal mortality even within narrow birth weight strata. It was capable of separating patients into groups with 2 to 20 times higher mortality risk. It also correlated highly with other indicators of severity including nursing workload (r = .59), physician estimates of mortality risk (r = .65), and length of stay (R^2 = .59).

Subsequently a Simplified Neonatal Illness Severity and Mortality Risk Score (Table 7) (SNAP-II and SNAPPE-II) was developed in thirty neonatal intensive care units in Canada, California, and New England during the mid 1990s. SNAP-II and SNAPPE-II are empirically validated illness severity and mortality risk scores for newborn[34].

The primary outcome was in-hospital mortality. Patients moribund at birth or discharged to normal newborn care in <24 hours were excluded. Starting with the 34 data elements of the Score for Neonatal Acute Physiology (SNAP), a logistic model for in-hospital mortality was designed using 10,819 randomly selected Canadian cases. SNAP-II includes 6 physiologic items; to this were added points for birth weight, low Apgar score, and small for gestational age to create a 9-item SNAP-Perinatal Extension-II (SNAPPE-II). SNAPPE-II was validated on the remaining 14,610 cases and optimized the calibration.

The discrimination of deaths from survivors by SNAPPE-II was excellent[34], with ROC areas ranging from 0.84 to 0.92 in the various populations and subgroups, with an overall performance of .91 ± 0.01, the weakest performance being for infants with birth weights <1500 g. For comparison, ROC areas for birth weight alone were .78 ± 0.01overall, .78 ± 0.01 for infants <1500 g, and .51 ± 0.02 for infants >1500 g.

Combining only birth weight, SGA, and low Apgar score (without SNAPII) showed ROC areas of .84, .81, and.76 for the 3 birth weight groups, respectively. The substantially higher discrimination by SNAPPE-II indicated the value of the additional physiologic derangement information.

Prognostic (predictive) scores

Paediatric Trauma Score (PTS)

The Pediatric Trauma Score (Table 8) is a trauma scoring tool for use in evaluating the severity of injury. The PTS[35] adjusts its scoring areas to account for the physiological and anatomical differences unique to the pediatric patient in turn more accurately identifying the critical patient. The PTS also allows for data gathering efforts which will be especially important with the development of a country wide trauma system and trauma registry. Locally the PTS can provide data on injury patterns for a geographical location which can be used to develop injury prevention programs and continuing education programs for EMS personnel.

The PTS consists of six parameters which are common determinants of the clinical condition in the injured child. During the initial assessment of the injured child each parameter is assessed and given a numeric score based upon its three associated variables: +2 (no injury or non-life threatening), +1 (minor injury or potentially life threatening), or -1 (life-threatening). Totals can range from a +12 to a -6 with the range of <8-9 being the critical break point for transport to a comprehensive pediatric trauma care facility.

Critically ill patients are typically characterized by disturbances of body homeostasis. Both in adults and children, these disturbances can be estimated by measuring how apart one or many physiologic variables are from the normal range. Composite scores are constructed with such variables.

Meningococcal Septic Shock Score MSSS

A sensitive scoring method for identification of children with presumed meningococcal septic shock (PMSS) (Table 9) at risk of death at admission to the pediatric intensive care unit was developed in 2001[38].The investigators wanted to develop a mortality prediction tool for PMSS and compare its performance with three other prognostic systems: 1) a generic mortality prediction tool, the Pediatric Index of Mortality (PIM), 2) the Glasgow Meningococcal Septicemia Prognostic Score (GMSPS) and 3) the Malley score. The study design was a multi-center retrospective cohort

The study involved all children aged between 1 month and 14 years with a confirmed or presumed diagnosis of meningococcal septic shock admitted to 14 PICU's of tertiary level hospitals in Spain.

The worst value of each variable during the first 2 hours in the PICU was selected for the analysis. 30 prognostic variables were tested: (demographic and clinical)[39]. Also included were 2 therapeutic variables: use of mechanical ventilation and refractory hypotension. The use of steroids as a predictor was not evaluated. The logistic regression identified seven independent predictors of death at admission: cyanosis, coma (GCS < 8), refractory hypotension, oliguria, WBC < 4000/mm3, PTT > 150% of control and base excess > -10mmol/l.

The primary outcome measure was hospital mortality, defined as death occurring before hospital discharge. A higher score on the model predicted a higher probability of death. The other three scores yielded lower ROC areas and this was statistically significant for the differences between the new score and the Malley and GMSPS ROC curves.

The 95% confidence intervals for the estimates of risk of death are not stated in the article. However, by converting the % mortality estimate in the three groups into proportions and using the raw numbers given in the article it is possible to calculate the 95% CI. Of note, the confidence intervals in the three risk groups do not overlap.

However, the new ICUs were still within the same geographic locale (Spain) and hence share important population characteristics. The prediction tool may work differently in a different population, e.g., USA, and so it is important to validate a tool in

several different clinical settings especially if the probabilities of death are different. Ideally, a validation of the cutoff values of the predictor variables (the prediction model) should be a) prospective and b) in a new population and c) with a different prevalence and spectrum of the underlying disease. The current study does not meet the first criterion as it was retrospective but does meet the second criterion fully. It meets the third criterion partially as the validation sample had an overall different case mix and study period than the development sample. Specifically, there were significantly more cases of sero group C, female sex and a shorter time to admission to the PICU from the time of appearance of petechiae.

The new score differs from a generic prediction tool like the PRISM and PRISM II in that it is derived from a more homogenous group of patients ("customized probability model"). Neither the new scoring system nor the comparison models were able to predict a risk group with 100% mortality. Information from the model about the child's average risk of dying, may be useful to surrogate decision makers like parents and to treating clinicians in making informed decisions about the risks/benefits of aggressive/expensive therapies.

Paediatric risk of mortality score (PRISM)

The Paediatric Risk of Mortality[18] (PRISM) score was developed from the Physiologic Stability Index (1986)[21] by Pollack et al. (it is named PRISM II score by some intensivists) to reduce the number of physiologic variables required for pediatric ICU (PICU) mortality risk assessment and to obtain an objective weighting of the remaining variables. Univariate and multivariate statistical techniques were applied to admission day PSI data (1,415 patients, 116 deaths) from four PICUs in the United States. The resulting PRISM score consisted of 14 routinely measured, physiologic variables, and 23 variable ranges. The performance of a logistic function estimating PICU mortality risk from the PRISM score, age, and operative status was tested in a different sample from six PICUs (1,227 patients, 105 deaths), each PICU separately, and in diagnostic groups using chi-square goodness-of-fit tests and receiver operating characteristic (ROC) analysis. In all groups, the number and distribution of survivors and

non survivors in adjacent mortality risk intervals were accurately predicted. ROC analysis also demonstrated excellent predictor performance (area index = 0.92 +/- 0.02).

PRISM is accurate and widely accepted, but many units do not use it routinely because it is difficult to collect the large amount of information needed to calculate it. PRISM scores should be used in critically ill neonates, infants, children, or adolescents, not in premature infants or in adults The variables used by PRISM that are not used by PIM are diastolic blood pressure, heart rate, respiratory rate, arterial partial pressure of carbon dioxide, the Glasgow Coma Score prothrombin time, serum bilirubin, serum potassium, serum calcium, blood glucose and plasma bicarbonate.

Updated Paediatric Risk of Mortality score (PRISM III)

The relationship between physiologic status and mortality risk was reevaluated as new treatment protocols, therapeutic interventions, and monitoring strategies are introduced, and as patient populations changed necessitated the development of a third-generation pediatric physiology-based score. In addition, since minimizing the time period for assessing mortality risk is advantageous for evaluating pediatric ICU quality, a 12-hr prediction model as well as a 24-hr prediction model was developed. The PRISM III was developed in 32 pediatric intensive care units representing about 10% of all pediatric ICUs in the United States[20], between 1993 and 1994. Consecutive admissions at each site were included until at least 11 deaths per site occurred.

Exclusion criteria included a) admissions for recovery from procedures normally cared for in other hospital locations; b) patients staying in the ICU less than 2 hrs; c) patients transferred from the study pediatric ICU to another ICU because their outcome could not be clearly credited to either ICU; and d) patients admitted in a state of continuous cardiopulmonary resuscitation who never achieved stable vital signs for at least 2 hrs.

An attempt was made at maximizing the predictive performance while keeping the number of variables and their ranges to a minimum, using variables that are readily available and clearly definable while maintaining the assumptions inherent in the

Physiologic Stability Index and PRISM that unmeasured variables were assumed to be normal. Therapeutic variables that may be unduly influenced by practice patterns were avoided.

Physiologic data included the most abnormal values from the first 12 and the second 12 hrs of ICU stay. Outcomes and descriptive data were also collected. Physiologic variables in which normal values change with age, were stratified by age (neonate, infant, child, adolescent). The data consisted of the following: systolic and diastolic blood pressures; heart rate; respiratory rate; temperature (oral, axillary, or core); coma status; pupillary reactions; pupillary size and equality; concentrations of sodium, potassium, total CO_2, bicarbonate, total and direct bilirubin, total and ionized calcium, glucose, blood urea nitrogen, creatinine, and albumin; hemoglobin; white blood cell count; platelet count; prothrombin and partial thromboplastin times; pH and PCO_2 (arterial, venous, or capillary); and PaO_2 with a simultaneous FIO_2. Whole blood as well as serum/plasma measurements of sodium, potassium, and glucose were also collected. For variables where both high and low abnormalities may reflect increased mortality risk, both the high and the low values were collected. Thus, both high and low values of the same physiologic variable could contribute to severity of illness. Physiologic data accumulated during the pre terminal period in patients dying within the first 24 hrs of pediatric ICU care were not included in the study when death was obvious.

When several variables overlapped significantly in the assessment of physiologic dysfunction, they were combined into a composite variable. For example, pH and total CO_2 were combined into a variable representing acidosis.

The database was randomly split into development (90%) and validation (10%) sets. Variables and their ranges were chosen by computing the risk of death (odds ratios) relative to the midrange of survivors for each physiologic variable. Data were collected on 11,165 admissions (543 deaths). The PRISM III score has 17 physiologic variables subdivided into 26 ranges. The variables most predictive of mortality were minimum systolic blood pressure, abnormal pupillary reflexes, and stupor/coma. Other risk factors, including two acute and two chronic diagnoses, and four additional risk factors, were

used in the final predictors. Variables in the original PRISM that are not included in PRISM III are diastolic blood pressure, respiratory rate, PaO_2 /FIO_2, and bilirubin and calcium concentrations[20]. Variables which are included in PRISM III but not in PRISM are temperature, pH, PaO_2, creatinine concentration, blood urea nitrogen concentration, white blood cell count, and platelet count.

The PRISM III score and the additional risk factors were applied to the first 12 hrs of stay (PRISM III-12) and the first 24 hrs of stay (PRISM III-24). There were no significant calibration errors. The area under the receiver operating curve and Flora's z-statistic indicated excellent discrimination and accuracy (area under the receiver operating curve for PRISM III-12 (.947 +/- 0.007); PRISM III-24 (0.958 +/- 0.006).

Generally, the PRISM III-24 performed better than the PRISM III-12 models. Excellent goodness-of-fit was also found for patient groups stratified by age, and by diagnosis. PRISM III-24 was compared with the original PRISM, area under the receiver operating curve improved by 3.9% (PRISM 0.914; PRISM III 0.950, p less than .0001).

PRISM III resulted in several improvements over the original PRISM. Reassessment of physiologic variables and their ranges, better age adjustment for selected variables, and additional risk factors resulted in a mortality risk model that was more accurate and discriminates better. First, the physiologic variables and their ranges were reevaluated thereby eliminating some ranges that did not contribute significantly to mortality risk (e.g., high systolic blood pressure). Although some physiologic variables have been eliminated and others added the variables with the greatest importance in outcome prediction were the same in both scores: low systolic blood pressure, altered mental status, and abnormal pupillary reflexes.

While age was included as an explicit variable in the original PRISM score, it was included in the PRISM III score in a logically and clinically more convincing form by using appropriate age-adjusted physiologic variable ranges.

The relationship between physiologic status, as measured by PRISM III, and outcomes has been calibrated to a contemporary, diverse, well-defined, large reference

sample.. These units encompass a wide diversity of organizational structure and patient mixes. This diversity makes the sample sufficiently representative for most units, enabling PRISM III to be used in the comparative assessment of pediatric ICU outcomes in essentially all pediatric ICUs.

PRISM reduced the number of physiologic variables to 14 and their ranges to 34. The total number of ranges in PRISM III was reduced. PRISM III-24 incorporates the most information over the longest time period. The use of the PRISM III-12 model is appealing for quality assessments since, by shortening data acquisition time, it better separates the observation from the treatment period, while the PRISM III-24 model is more accurate for individual patient mortality risk assessments. The large number of diverse ICUs in the database indicates PRISM III was more likely to be representative of United States units. A serious problem with 12 or 24-h scores was that they were affected by treatment given after admission to intensive care, so that they are not valid instruments for comparing the quality of care between different units, or within a single unit over time. Children admitted to a good PICU who recovered were found to have lower PRISM scores than similar Children admitted to a bad PICU who are mismanaged in the first12–24 h, and the bad unit‘s high mortality rate will be incorrectly attributed to its having sicker patients.

Development of the Paediatric index of mortality scores (PIM and PIM2)

Paediatric index of mortality (PIM)

Towards the end of the 20th century a need was felt for models that predict the risk of mortality in children in intensive care which would allow evaluation of the effectiveness and efficiency of paediatric intensive care. Models to compare the standard of care between units and within units over time by adjusting for differences in severity of illness and diagnosis were required. Comparing different systems of organising intensive care was another need at that point. Estimating mortality risk was also considered an important component of comparing groups of patients in research trials. One such score was the PIM score.

The initial score was developed using three prospective cohort studies from the year 1988 to 1995[22]. A fourth cohort study, from 1994 to 1996, collected information

from consecutive admissions to seven paediatric intensive care units in Australia and one in Britain. The development of PIM began in 1988, when information was collected about 678 consecutive admissions over 6 months to the PICU at the Royal Children‘s Hospital, Melbourne. The variables collected were the 34 PSI (physiological stability index) variables[21], along with mean arterial pressure, ventilator peak inspiratory pressure (PIP), ventilator positive end-expiratory pressure (PEEP), motor response to pain, immature neutrophil count, total neutrophil count, base excess, and rectal temperature. The worst value of each variable in the first 24 hours after admission as recorded for all 678 patients, and the admission values were also recorded for the last 230 patients.

The second stage[22] of the study began in 1990, when 814 consecutive admissions to a PICU at Melbourne were studied. Information was collected at the time of admission and over the first 24 h in PICU about age, gestational age, pupil reaction to light, and motor response to pain, base excess, mean arterial pressure, and respiratory rate, arterial carbon dioxide tension (PaCO2), PIP and PEEP.

In the third stage[22] of the study, from February 1994 to March 1995, 1412 consecutive admissions to the PICU at Melbourne were studied. Information was collected at the time of admission to PICU and during the first 24 h about all PRISM variables plus information about sex, time in hospital before admission to PICU, need for mechanical ventilation, diagnosis, the presence of a right-to-left cardiac shunt, estimated fractional inspired oxygen concentration (FIO2) in unintubated patients, weight, mean blood pressure, each pupil‘s size and reaction to light, PIP, PEEP, PaCO2, base excess, and plasma sodium. The above parameters were then evaluated for an association with mortality. The continuous variables were appropriately transformed and tested for an association with mortality. With the exception of age and time in hospital before admission to intensive care (lead time), variables that were not associated with mortality on Univariate testing ($p>0.1$) were excluded from further analysis. A preliminary model was developed. In the fourth stage of the study, during the period 1994–1996, information about the variables in the preliminary model (plus plasma sodium and prothrombin time) was collected from consecutive admissions less than 16 years of age to four PICUs in Australia (the learning sample) and one PICU in Britain and three PICUs in Australia (the test sample). Each unit collected data from enough consecutive

admissions to include at least 20 deaths. As a check on the accuracy of data collection, a sample of the data was collected in duplicate. The information from the above was used as a learning sample to determine the regression coefficients of a logistic model. The fit of the model developed on the learning sample was then tested on children admitted to the PICUs at Birmingham, Brisbane, and Adelaide. Calibration of the model, to evaluate how well the model classifies subjects into low, medium and high risk categories, was done. Inspection of the number of observed and expected deaths and survivors in five groups with <1%, 1–4%, 5–14%, 15–29% and 30% or more predicted mortality was also carried out. Discrimination estimates how well the model distinguishes between patients who lived and patients who died. Once the model was found fit in both the development and the validation groups, logistic regression coefficients were re-estimated using the entire sample. The first, second and third parts of the study were used to determine the variables that were included in the fourth part of the study. A total of 5695 children were in the fourth part of the study, and 278 of them died. Six of the 278 deaths occurred within 24 h of discharge from intensive care. No patient was lost to follow-up[22].

Physiological variables that were not measured were considered to be normal. Values for systolic blood pressure, base excess were transformed using statistical methods[24,25]. The predictive power of pupil size, inequality in size, and their reaction to light was obtained.

Each child was allocated one of 214 different diagnoses; nine diagnoses were associated with an increased risk of mortality, even when information from the other PIM variables was taken into account.

Prediction was not improved by having a different coefficient for each of the nine diagnoses, by the inclusion of interaction terms in the model, or by including the square of the transformed base excess variable, the square of FIO2/PaO2, or the logarithm of transformed systolic blood pressure. Variables that did not predict death were serum bilirubin, pulse rate, central venous pressure, haemoglobin, the presence of convulsions, left atrial pressure, and days in hospital before admission to intensive care (the lead time). Variables such as the prothrombin time and the serum sodium value were statistically significant when added to the final model, but were excluded because they had little effect on the fit of the model. So the final parameters as chosen by the PIM score were

Pupils fixed to light, Specified diagnosis, Elective admission, Mechanical ventilation, absolute (SBP-120) mmHg, absolute (base excess), 100´FiO2/PaO2 ($mmHg^{-1}$).

Some of the other salient findings of the study to device the PIM score include the fact that neonates had a higher mortality than older children, but inclusion of age did not improve the prediction of the model. Correction for age did not improve the predictive power of systolic or mean blood pressure, pulse rate, or respiratory rate. None of the PICUs had a significant effect when they were included in the model as dummy variables.

The final model was estimated using the entire sample from the fourth part of the study. The deciles of risk goodness-of-fit test gave p=0.37 and the areas under the ROC plot was found to be consistently above 0.80[22]. The model described their risk of mortality for babies less than one month well. A training model was developed on one group of intensive care units and was applied to another group of unit so as to provide a more stringent test than just randomizing individual patients to the training or test set[26].

Death was chosen as the dependent variable in this model, rather than death in hospital or death within one month of admission into a PICU. Death in PICU is the mortality outcome that is of most practical interest to paediatric intensivists and was the outcome used by PRISM[17]. Multiple admissions for an individual were included in the study because the model was used on data that includes children who are admitted several times.

Mortality prediction models, such as the PIM are developed by finding variables that predict the probability of death in groups of children. This model is then often used as a measure of severity of illness, which assumes that children with a high risk of death are sicker than children with a low risk of death. This assumption may be true for many types of PICU patients, but not all. An example quoted by the authors of the PIM include children with epiglottitis or severe croup are very likely to die without intensive care, but they have a very low mortality if they are properly managed – so the PIM would give these children a low score despite the fact that they are very ill. Although the PIM may provide a fairly good description of groups of patients, it may not be accurate enough to be used to make decisions about the management of individual patients.

The PIM model was simple enough for it to be widely used in paediatric intensive care – it required the collection of only eight variables at the time of admission to intensive care, while having a good predictive power. The PIM was developed in dedicated PICUs where there are high levels of consultant input, senior resident staff and trained PICU nurses.

Paediatric index of mortality 2 score (PIM2)

Outcome in terms of mortality is influenced by new treatments and new management approaches. Changes in referral practices, systems of providing intensive care, attitudes to the indications for commencing and discontinuing life support lead to changes in thresholds for admission to intensive care and alterations in the relationship between disease and outcome. Further, as experience, and therefore the quantity of data, expands it is possible to use a larger and more diverse patient population to develop mortality prediction models.

The objective behind development of the PIM2[27] (Table 10) was to revise the Paediatric Index of Mortality (PIM) to adjust for improvement in the outcome of paediatric intensive care, through an international multi centre observational study carried out in twelve specialist paediatric intensive care units and two combined adult and paediatric units in Australia, New Zealand and the United Kingdom. 20,787 patient admissions of children, less than 16 years, admitted during the study period were included. Again there were no interventions[27]. A revised model was developed by forward and backward logistic regression. Variable selection was based on the effect of including or dropping variables on discrimination and fit. The addition of three variables, all derived from the main reason for ICU admission, improved the fit across diagnostic groups. Data from seven units were used to derive a learning model that was tested using data from seven other units. The model fitted the test data well and discriminated between death and survival well (area under the receiver operating characteristic (ROC) plot 0.90 (0.89-0.92)). The final PIM2 model, derived from the entire sample of 19,638 survivors and 1,104 children who died, also fitted and discriminated well [^{2} 11.56, p=0.17; area 0.90 (0.890.91)[27].

Ten Australian and New Zealand intensive care units, and 4 British units[28] collected uniform paediatric data commencing on January 1, 1997. The data included the PIM variables, demographic variables, the principal ICU diagnosis (defined as the main reason for ICU admission) and ICU outcome (died in ICU, discharged or transferred to another ICU). Data from the above were combined to develop and validate a revised model, PIM2. All patients admitted consecutively during the period of study were included. Patients 16 years or older were excluded, as were patients transferred to other ICUs, because these patients could not be appropriately classified as ICU survivors or deaths.

The first step in revising the model was to examine the ratio of observed deaths to deaths predicted by PIM in the entire population, and when patients were grouped by mortality risk, diagnosis, diagnostic group, intensive care unit, and age[27]. The aim was to identify patient groups where PIM either over-predicted or under-predicted mortality. Individual variables were examined for association with mortality using appropriate statistical methods. When appropriate, transformation was used to improve the relationship between a variable and mortality, followed by forward and backward logistic regression to test each of the original variables and potential additional or substitute variables.

To test the revised model, the population was divided into a learning and test sample by randomly selecting units, stratified by size of unit and country. The logistic regression model developed in the learning sample was evaluated in the test sample by calculating the area under the receiver operating characteristic plot to assess discrimination between death and survival[29]. Calibration across deciles of risk was evaluated using the Hosmer-Lemeshow goodness-of-fit test. To examine the fit of the model in more detail, tables were constructed to assess calibration across risk, age, and diagnostic group by visual inspection of the number of observed and expected deaths.

The reproducibility of data collection in each unit in was assessed by repeating the data collection for 50 randomly selected patients stratified by mortality risk. Data quality was assessed by comparing the probability of death predicted by PIM in the two data sets. The results of this study revealed 1,104 deaths, giving a mortality rate of 5.3%. Logistic regression was used to generate new coefficients for the original PIM variables.

This re-calibrated, first generation model was then tested. Discrimination was adequate (area under the curve in a ROC was 0.88 (CI 0.87-0.89))[27], however, calibration across diagnostic groups was poor in two groups: respiratory illness and non-cardiac post-operative patients (observed: expected deaths, 160:212.8 and 48:82.0, respectively). The PIM variable "Specific Diagnosis" included nine diagnoses associated with increased risk of death. The diagnoses –in hospital cardiac arrest and liver failure were associated with increased risk of death and five common diagnoses (asthma, Bronchiolitis, croup, obstructive sleep apnoea, diabetic keto-acidosis) were associated with reduced risk. The mortality of patients admitted primarily for post-operative recovery was better than predicted by PIM for all surgical groups except for patients admitted following cardiac bypass. "Specific Diagnosis" seen in the PIM was replaced by "High Risk Diagnosis" and "Low Risk Diagnosis" in the PIM 2. Pupillary reactions as defined in PIM remained a significant predictor both in Univariate and multivariate analysis. No change was made to the four physiological variables from the original model.

In the test sample the new model discriminated well between death and survival [Area under the curve (Az) ROC 0.90 (0.89-0.92)] and calibrated across deciles of risk well (goodness of fit ^{2} 8.14, 8df, p=0.42). The final PIM2 model estimated from the entire sample also discriminated and calibrated well (Az ROC 0.90 (0.890.91); goodness-of-fit test ^{2} 11.56, 8df, p=0.17)[27]. The performance in respiratory illness and non-cardiac post-operative patients was improved in the revised model.

PIM2 was derived from a larger, more recent and more diverse data set than the one used for the first version of PIM[27]. Three variables, all derived from the main reason for ICU admission, had been added to the model (admitted for recovery from surgery or a procedure, admitted following cardiac bypass and low risk diagnosis). Changes have been made to the variable "High Risk Diagnosis": the criteria for cardiac arrest had changed, liver failure was been included and IQ below 35 omitted.

On application of the original PIM model to the second data set, the overall standardized mortality ratio (SMR) was 0.86 (0.81-0.90) with similar values encountered in all units. However 14% of the children predicted to die using 1994-1995 standards survived in 1997-1999[27]. The explanation for this improvement was not known. It was considered that that incremental gain had been achieved by improved application of old

therapies. Critically ill children were probably being recognized and referred earlier with good effect. It was not possible to test these or other hypotheses on the current data and the explanation for the apparent improvement in management is unknown. The standard of care set by the original PIM was high[22] compared to PRISM and PRISMIII with the standard set by PIM2 is even higher[27].

The same methods were used in the development of both models. Variables were included only if they improved the discrimination or calibration of the model. To test the new model, the data were split into two groups of units. Coefficients derived on the learning sample demonstrated good performance in the test sample. The coefficients derived from the entire sample were used in the final model.

As well as calibrating across mortality risk, it was considered important that intensive care prediction models calibrate across diagnostic groups. If the model was found to over- or under-predict mortality in a large group of patients, then the overall performance of the unit assessed by the model would be influenced by the proportion of patients admitted in this category. This was found to result in biasing the estimated SMR but also exposed the added danger that units would dismiss potentially important results because they attribute an unexpected finding to the mix of patients rather than the standard of care. The first version of PIM over-predicts death in non-cardiac post-operative patients and, to a lesser extent, respiratory patients. This trend was evident in the original study; however, although there were more than 5000 patients in that study, there were only six deaths in non-cardiac post-operative patients. In the study to develop PIM 2 there were 48 deaths in this diagnostic group and the tendency for the original PIM to over-predict death was confirmed. The addition of variables that identify diagnoses with a low risk of mortality has improved the performance of PIM2 in non-cardiac post-operative patients and respiratory patients. The diagnoses included in the variables "High Risk Diagnosis" and "Low Risk Diagnosis" represent conditions where the physiological and demographic PIM2 variables either under-estimate or over-estimate the risk of death. Diagnoses associated with high or low risk of death where the risk was accurately predicted by the physiological and demographic PIM2 variables were not included in the specific high or low risk diagnoses.

Examination of the model performance in specific diagnoses resulted in other changes to the model. The risk-adjusted outcome for cardiac arrest preceding ICU admission was similar for in-hospital and out-of-hospital cardiac arrest; therefore PIM2 does not restrict cardiac arrest preceding ICU admission to out-of-hospital cardiac arrest. Liver failure (acute or chronic) as the main reason for ICU admission has been added to the list of high risk diagnoses. The specific diagnosis "IQ below 35" was removed, primarily because it proved difficult to code reproducibly, particularly in young children. Omitting this diagnosis from the model altered the area under the ROC plot by less than 0.1%[27]. A major advantage of using admission data to estimate the mortality risk was that the model was not biased by the quality of treatment after admission. In this respect PIM is preferred to models that use data collected during the first 12-24 h after admission. A potential criticism of PIM, however, was that one of the variables, mechanical ventilation during the first hour, was also susceptible to bias resulting from different intervention thresholds. In the PIM2 study the percentage of patients intubated during their ICU stay varied between units from 25 to 93%.

Mechanical ventilation during the first hour could be considered a simple way of accounting for variation in admission thresholds and weighting the model for patients that require life support. Ideally, mortality prediction models should not be influenced by treatment. Omission of the ventilation variable from the model, resulted in a drop in the the area under the ROC plot from 0.90 to 0.88[27].

Accurate data collection is critically important. There is an increase in complexity with the number of variables in the model increased from seven to ten in PIM2, the additional variables being derived from the principal ICU diagnosis, which is an important variable for all ICU patients. When data is collected as given in the PIM 2 guidelines rather than deriving these codes electronically from a database containing the diagnosis, it allows a consistency of coding and also enables electronic checking against the diagnosis to be used as a technique to verify the data. It was thought not advisable for data to be collected by large numbers of doctors and nurses as an addendum to their routine clinical work. In this study, the 95% confidence intervals for the bias in risk of death estimation included 1.0, suggesting that data collection errors were not significantly affecting prediction.

When the score did not predict the correct number of deaths in a unit, it was considered that the standard of care respective unit was better or worse than the standard in the units that developed PIM2 in 1997-1999, or that the characteristics or diagnoses of patients in the unit were substantially different from the population in the original study. Changing the coefficients in the model to ensure better outcome prediction defeats one of the main purposes of the model, which is to allow units to compare their performance with that of the Australasian and UK units that developed it. PIM2 was applied only to groups of patients and not allowed to describe or influence the management of individual patients.

Details on instructions for use of the PIM2 score are given in appendix 1.

Results of various studies world wide using the PIM are discussed in the tables 11 and 12.

Differences between the PIM and the PRISM

Shann et al[22] developed the PIM score as an alternative to the PRISM III [20] for comparing observed with predicted mortality among PICUs. Any score should be valid, reliable, accurate, easier to use and cost-effective. Whether mortality prediction scores are a valid measure of the quality-of-care delivered by a PICU, is an important question.

The PIM has a simple job to do: to predict death for patients who die and survival for those who live. The evaluation of the ability of a score to discriminate among these two populations is described by the area under the receiver operator characteristic (ROC) curve. This curve is a plot of true-positive versus false-positive predictions. For the PIM this area was 0.90. PIM discriminates well in a new populations of PICUs not used in the derivation of the initial score, therefore validating the PIM. In comparison, PRISM III [2] also discriminated well (ROC area=0.94) in the separate validation sample comprised of 10% of patients from the 32 PICUs used in its development.

PIM uses data present upon admission rather than use the worst values during the first 12 or 24 h after admission, as is done when using the PRISM III. Patients with lower predicted mortality scores upon admission who receive “bad” care and deteriorate within the first 24 h will be counted as unexpected deaths if admission scores are used, but will be counted as expected deaths if the most abnormal values in 24 hours are used.

However values reflecting the physiological status present upon admission could reflect a transient state resulting from interventions during transport or in the operating room. To date, there is no consensus has been reached as to which approach is better.

Reliability is the ability of a score to predict mortality accurately in various subgroups of patients. The PIM is well calibrated in each decile of mortality risk. However, when patients are categorized by diagnostic groups (for example in the group comprising postoperative non-cardiac patients) the area under the ROC curve was significant. The PIM discriminates quite well for trauma patients (ROC area=0.94), but not as well in cardiac patients (ROC area=0.83)[27].

Another test of reliability is the ability of various raters to derive the same score when assessing a patient at the same point in time. The PIM and the PRISM III perform well on this evaluation criterion.

In a subset of 1182 children from a single Australian PICU, PRISM predicted 118.6 deaths, PIM predicted 71.6 deaths, and there were 78 actual deaths[23]. Over prediction of mortality by the PRISM score could be because the centralized care delivered by this large tertiary centre is superior. However poor performance of mortality prediction scores in countries where they were not developed has been shown in other studies[36].

The PIM is a score derived in an Australian population made up of large centralized PICUs. Patients cared for in small community hospital ICUs may differ in ways that are not adjusted for by the score. The PRISM III was developed using randomly selected PICUs while ensuring that a variety of ICU characteristics were present in the development sample.

The PIM outperforms the PRISM III in its ease of use. The PIM requires collection of only 8 variables upon admission. The PRISM III requires collection of the most abnormal (highest and lowest) values of 17 physiologic variables during the first 24 h after admission plus 6 additional risk factors. In recent times data collection has been made easier with automated data collection systems. The coefficients for the PIM are freely available, whereas use of the PRISM III requires payment of a fee. With each

update of the PRISM score, the coefficients have been readjusted to reflect differences in patient outcomes that have developed as processes of care and patient characteristics change over time.

Both the PIM and the PRISM III claim that comparison of observed versus expected mortality among a group of PICUs is a beneficial technique for rating PICU performance. Since mortality is a relatively rarer outcome, in both the PIM and the PRISM III patient cohorts, the majority of care being delivered may not be assessed by mortality prediction scores. These scores do not predict other more common outcomes such as length of stay or functional status very well. Both the PIM and the PRISM III are not accurate for individual prognostication.

The PRISM was found to predict 66% more deaths than the PIM in an Australian PICU possible due to highly centralized intensive care in Australia. The variables used by PIM that are not used by PRISM are the presence of a specified diagnosis, use of mechanical ventilation and the plasma base excess. The PRISM III [20] had an area under the ROC plot of 0.94 for both the 12-h and 24-h models – but may be considered to be even more complicated than the original version[19]. An annual license fee was collected for use of the PRISM scores.

Treatment given just before admission to intensive care is likely to affect admission scores (such as PIM) more than 24-h scores. For example, in a patient with shock, appropriate administration of fluid and sympathomimetics may increase blood pressure and restore the base excess to normal, which will affect the PIM score. However, if this treatment improves the patient‘s prognosis at the time of admission to intensive care, it alters the PIM score. It was suggested that patients with a given severity-of-illness score may have a higher mortality rate if they have been extensively treated before they are admitted to intensive care[37], a problem known as lead time bias, but time spent in hospital before admission to intensive care was not found to be statistically significant when added to the PIM model. PRISM uses data collected over the first 24 h after admission to intensive care (12 or 24 h for PRISM III), and many of the deaths occur during this time, so that the score may be diagnosing death rather than predicting it in some patients.

Therefore in conclusion on reviewing the literature the PIM has been proved to be an adequate indicator of the risk of mortality world wide in the subset involving critically ill children.

Materials and methods

Study design

The design of the study was an observational prospective cohort study

Study population

All children admitted into the Paediatric Intensive Care Unit (PICU) from 1st February 2007 to 31st September 2007 were included in the study. There was no intervention in this study. All consecutive children were included in the study and none were excluded due to any reason.

Study setting

This study was performed in Paediatric Intensive Care Unit (PICU) of Department of Child Health, Christian Medical College and Hospital, Vellore which is a multispecialty tertiary care hospital in South India. The paediatric wards have 133 general beds and cater services to General Paediatrics, Paediatric Oncology, Paediatric Nephrology, Paediatric Endocrinology, Developmental Paediatrics and Paediatric Surgery.

The PICU has eleven beds and five step down semi ICU beds. The PICU not only accepts children from all above mentioned Paediatric specialties but also from other specialties such as Hematology, Urology, Plastic surgery and Cardiology. Children are received directly from the Paediatric Emergency services or from General Paediatric wards. Paediatric surgical patients are admitted either from their ward or from operating theatres. Out born neonates and inborn neonates aged over 72 hours are also admitted to the PICU. How ever paediatric population requiring post operative care in the field of

cardiothoracic surgery and neurosurgery were managed separately under their respective ICUs.

The PICU is staffed by one senior consultant trained in intensive care, two junior consultants and five medical officers who work in shifts. There are four registered nursing staff per shift with the nurse: patient ratio of 1:3. This team also includes two respiratory therapists and one dialysis technician.

The intensive care unit can provide the following modalities of treatment: ventilatory support, inotropic support, cardiovascular monitoring, central venous access, invasive CVP and arterial monitoring and ETCO2 monitoring. Peritoneal dialysis was the predominant mode of renal support while this study was being done. Intermittent hemodialysis and hemofilteration are technically possible and were being increasingly utilized. Blood gas analysis was done in the Biochemistry department and was available 24 hours of the day.

Sample size

As per the statisticians advice a defined period of study was chosen before the onset of the study. This period was chosen as minimum of eight months. All children admitted under PICU for treatment during the period were included in the study.

Data collection

The PIM 2 variables were collected within one hour of admission into PICU as set out for PIM 2 score (refer appendix 1). All the data was collected by the primary investigator. Patients with multiple admissions were included and regarded as separate admissions. Demographic data, physiological data and the clinical diagnoses were entered at the time of admission. Informed consent was not taken as there was no

intervention in the study design. The admission diagnosis was entered and classified into various diagnostic groups consisting of CNS, respiratory, cardiovascular, gastrointestinal or liver, sepsis, multi-organ failure, hematological, poisoning, metabolic, renal, post-operative and others. The length of stay in PICU and final outcome (Death or transfer out of PICU) were entered later on. No extra tests were performed to meet the needs of this research since it considered non-collection as normal.

Statistical analysis

PIM 2 SCORE was calculated for each patient using the software published in the website of www_sfar_org-scores-pim2.htm which is a free website available online. All the above collected DATA and the PIM 2 score were entered into the Statistical Package for the Social Sciences (Windows version 15.0; SPSS Inc, Chicago [IL], United States) and the DATA was analyzed.

Demographic and physiological data were described using median and interquartile ranges (IQRs) because they were not normally distributed. The individual mortality risk was predicted by a logistic regression equation and the overall predicted risk of intensive care unit mortality was subsequently calculated.

We tested the fit of PIM2 mortality prediction model in two ways – its ability to discriminate and calibrate the mortality risk. Discrimination refers to the ability of the test to calculate a higher mortality probability among non-survivors than survivors across the whole group. Calibration signifies how well the test predicts both mortality and survival across subcategories.

Discrimination was assessed using the area under the receiver operating characteristic curve (ROC). Receiver operating characteristic (**ROC**) curve is a graphical

representation of the relationship between the sensitivity and the specificity of a particular test. This measure expresses how well the model distinguishes between patients who lived and those who died. The area under the curve (**AUC**) represents the overall accuracy of the test. The larger the area the better is the test. An area under the ROC curve of 0.75 or more is considered clinically useful, and in the design of mortality prediction there is a trade off between the simplicity of the model (easier data collection and quality) and enhanced area under the ROC plot. An area under the ROC plot of 0.75 means that a randomly selected non-survivor would have a higher PIM value than a randomly selected survivor 75% of the time; it does not mean that prediction of death is correct 75% of the time.

Calibration of PIM 2 score evaluates how well the model classifies patients into low, medium, and high risk categories. We evaluated this by examining a Hosmer–Lemeshow goodness of fit table where acceptable calibration is evidenced by a p value $\geq$ 0.10. It compares model performance (observed versus expected) at five different risk intervals across the deciles of risk (i.e., predicted mortality <1%, 1–4%, 5–14%, 15–29%, or 30% or more) to test whether the model is biased (i.e. performs differentially at the extremes of risk). A non significant value for this test suggests an absence of such bias.

We also calculated the Standardized Mortality Ratio (SMR) , which is the ratio of observed to expected death rates. The overall expected death rate is the sum of the probability of death for each admission. Observed deaths included all children who either died in the ICU or were discharged in a terminal condition at parental request. Confidence intervals (CIs) for the standardized mortality ratios were also calculated. An

SMR value less than one would imply good performance of our PICU and value greater than one would imply poor performance. On the other hand, an SMR less than 1 could be interpreted as an overestimation of mortality in the PICU by the PIM. If the SMR was equal to one or if the CIs included the value 1.0 then it would indicate that the outcome of our patients was as predicted by PIM-2, indicating that the particular PICU is performing as par with the Australian PICU where it was developed initially.

RESULTS

A total number of 755 children were admitted into the Paediatric Intensive Care Unit (PICU) during the study period from 1st February to 31st September 2007.

Demographic pattern of the study population

There were 435 (57.6%) boys and 320 (42.4%) girls in the study population with the Male: Female ratio of 1.4:1 as shown table 1.

Table1:- Gender distribution

Gender	No. of children	%
Male	**435**	**57.6**
Female	**320**	**42.4**
Total	**755**	**100**

Table 2:- Age wise distribution

Age group	No. of children	%
Less than 1 year	305	40
1-5 years	246	33
6-10 years	102	14
11-15 years	102	14
Total	**755**	**100**

The age wise distribution of the 755 children is shown in table 2. Majority of the PICU admissions were infants below 1 year of age contributing to 305(40%) of the study population. This was followed by the 1-5 year age group contributing 246 (33%) children and 6-10 year and 11-15 year age groups contributing 102 (14%) children each as mentioned in the Table 2. The median age of the study group was 12.00 months with an interquartile range of 4-72months.

Table 3 shows that among the infants, neonates made up 30% (92 0f 305) of the admission.

Table 3:- Age distribution of Infants

Age group	No of Children	%
Less than 1month	92	30
1-6 months	135	44
7- <12months	78	26
Total	305	100

Graphical representation of the study population revealed that the age distribution of the study population did not follow a normal distribution curve as shown in the Figure 4.

Distribution pattern of the admissions

Majority of children got admitted from Paediatric accident and emergency services making up 59.4% (449 of 755) followed by ward transfer- in making up 37.3% (281of 755). Transfer in for post operative monitoring from operative theatres constituted 2.8% (21 of 755) of the PICU admissions as shown in the table 4 and Figure 5.

Table 4: - Pattern of PICU admissions

Admissions from	No. of children	%
CASUALTY	449	59.4
WARD	281	37.3
OT	21	2.8
OPD	3	0.4
BMTU	1	0.1
Total	755	100

Table 5:- Distribution of admissions based on primary admitting units

Table 5 and Figure 6 show the distribution of the children according to the admitting unit. 82.2% (621 of 755) were from the three general Paediatric units, 5.6 %(42 of 755) from Paediatric oncology unit, 1.4% (10 of 755) from Paediatric nephro-logy unit, 0.5% (4 of 755) from Paediatric Endocrinology. Paediatric surgery patients made up 7.5% (56 of 755). Rests of the specialties contributed 2.8% (22 of 755) to PICU admission.

Disease pattern in PICU

Table 6 and Figure 7 shows that Sepsis syndrome top the list of admissions in PICU constituting 142 of 755 (18.8%) admissions. This was followed by central nervous system illness in 131 of 755 (17.4%), respiratory illness in 97 of 755 (12.8 %), GI and hepatic illness in 94 of 755(12.5%). Poisoning and envenomations contributed 70 of 755 (9.2%), malignancy by 52 of 755 (6.9 %), cardiovascular illness by 45 of 755 (6 %), Surgical illness by 31 of 755(4.1%), renal illness by 26 of 755 (3.4%), Hematological illness by 20 of 755 (2.6%). Miscellaneous illnesses contributed to 37of 755(6.3%).

Admitting Units	No. of children	%
CH I	173	23.0
CH II	224	29.6
CH III	224	29.6
PAED.SURGERY	56	07.5
PAED. ONCOLOGY	42	05.6
PAED. NEPHROLOGY	10	01.4
PAED.ENDOCRINOLOGY	4	00.5
HAEMATOLOGY	8	01.0
MISC	14	01.8
Total	755	100

Table : 6 Disease Pattern In PICU

Primary System Involved	Number	%
Sepsis/MODS	142	18.7
CNS	131	17.3
Resp	97	12.8
GI	94	12.4
Poisoning	70	9.2
Malignancy	52	6.8
CVS	45	5.9
Post op	31	4.1
Renal	26	3.4
Hemat.	20	2.6
Rheumat.	10	1.3
Endo.	9	1.1
Accidents	8	1
IEM	5	0.6
Immunodef.	3	0.3
Misc.	12	1.5
Total	**755**	**100**

Frequency of PIM2 score variables

Table 7 shows that 725 (96%) children were emergency admissions and only 30 (4%) were elective admissions. 66 (8.7%) children were admitted for Post- procedure monitoring. There was no single admission due to cardiac bypass. Among the other PIM variables, 73 of 755 (9.7%) had high risk diagnosis and 24 of 755 (3.2%) had low risk diagnosis. Rest of the 658 children could be coded neither as high risk nor low risk.

At admission 27 children (3.7%) had dilated fixed pupils, 254(33.6%) children required mechanical ventilation within one hour admission into PICU and 125(16.6%)

admissions came in shock. 235 (31%) children had Base excess samples results in the first hour of their admissions. Fio2/PaO2 was available for 29.7% (224 of 755).

Table -8 shows the distribution of the children with the high risk diagnoses. Liver failure tops the list of high risk category (19 of 73) which was followed by leukemia after the first induction (18 of 73). Table -8 shows the distribution of children with low risk diagnoses. Bronchial asthma (13 of 24) tops the list.

Duration of PICU stay

Figure-7 and Table-9 show the duration of PICU stay. Out of the 755 children admitted, 586 (78%) had PICU stay of < 72hours, of which 96 (13%) stayed less than 24 hours and 490 (65%) stayed between 24-72hours. The median duration of PICU stay was 48 hours with range of 1 – 75 days.

Table 9:- Duration of PICU stay

Duration of stay	No of Children	%
< 24 hours	96	13
24-72hours	490	65
4-7days	126	16
8-14days	34	5
>14 days	9	1
Total	**755**	**100**

Figure 8 and table 10 show the relationship between the duration of PICU stay and the percentage of mechanical ventilation. There was a positive relationship between ventilatory support received and the duration of PICU stay.

Table 10:- Duration of stay Vs Mechanical ventilation

	Mechanical ventilation		
Duration of PICU stay	**No (%)**	**Yes (%)**	**Total**
<24 hours	275 (76.2)	86(23.8)	361
24-72hours	158(70.3)	67(29.7)	225
72-hours -1 week	52(41.3)	74(58.7)	126
1-2 weeks	13(38.3)	21(61.7)	34
>2weeks	3(33.3)	6(66.7)	9
Total	**501(75.6)**	**254(24.3)**	**755**

Among the ventilated children, electively admitted children required lesser duration of ventilation compared to emergency admissions. On the extremes of duration of PICU stay, mortality seems to be comparatively increased than definite as seen in table 11.

Table 11: Duration of stay Vs mortality

Duration of PICU stay	**Survival**	**Mortality**
<24hours	64(67%)	32(33%)
24-72hours	386(79%)	104(21%)
72hours -1week	95(75%)	31(25%)
1-2weeks	22(65%)	12(35%)
>2weeks	4(44%)	5(56%)
Total	**571**	**184**

Table 12:-*Outcome in PICU*

Outcome	No. of Children	%
Survival	572	75.6
Death	104	13.8
Discharged at request	80	10.5
Total	**755**	**100**

This table shows the outcome of the 755 children admitted to PICU. 184 (24.3%) children had an adverse outcome, of which death comprised of 104 (13.8%) and discharged against medical advice or by parental request 80(10.5%). Since these children were discharged in a terminal condition for the purpose of analysis we have considered them as mortality. The overall mortality was 24.3%.

Table 13 and Figure 10 show the distribution of diseases among the children who succumbed. Out of the 184 children, 62(33.6%) had sepsis with multiorgan dysfunction group followed by neurological illness in 25 (13.6%). Gastrointestinal- hepatic illness and malignancy constituted 19 (10.3%) each and respiratory illness in 16 (8.7%). Only one child with surgical problem had died.

Table 13 :- Disease status and the incidence of mortality

Disease	No of children	No of deaths	% of total mortality	% of mortality individual disease group
Sepsis / MODS	142	62	33.7	44
Accident and injury	8	1	0.5	12.5
CVS	45	13	7.1	29
Resp	97	16	8.7	16.5
CNS	131	25	13.6	19.0
Renal	26	1	0.5	3.8
Haemat	20	11	6	55
GI/Liver	94	19	10.3	20
Endo	9	2	1.1	22
Poisoning	70	5	2.7	7.1
Malignancy	52	19	10.3	36.5
IEM	5	2	1.1	40
Surgical	31	1	0.5	3.2
Misc	12	2	1.1	16.6
Immunodeficiency	3	2	1.1	66.7
Rheumat	10	3	1.6	30.0
Total	**755**	**184**	**100**	

Table- 13 also shows the deaths in the different diagnostic groups. Out of the 142 children admitted with sepsis and MODS 62(44%) had succumbed. Out of the 20 children with hematological disorders 11(55%) and among the 3 children with immunodeficiency diseases 2 (66%) had died. 19 out of the 52 children with malignancy also had an adverse outcome. However, hematological illness contributed only 6% of the overall mortality. Cardiovascular, respiratory and CNS disorders showed a mortality of 29%, 16.5% and 19.0% respectively. Out of 70 children with poisoning and envenomations 5(7.1%) had an adverse outcome. Lowest mortality was seen among the post-op surgical patients.

The performance of PIM2 score

During the study period 184 (24.3%) children out of 755 patients died. The overall expected deaths as predicted by PIM score was 183.9 with an area under the receiver operating characteristic curve of 0.839 with a 95% CI of 0.804 and 0.874 which was statistically significant (Figure-11).

Table 14: Calibration of PIM2 SCORE model by Hosmer Lemeshow goodness of Fit test

Outcome		**Survival**			**Mortality**	
Decile	**Observed**	**Expected**	**O:E ratio**	**Observed**	**Expected**	**O:E ratio**
1	65	57	1.14	0	8	0
2	52	49.8	1.04	5	7.2	0.69
3	44	44.5	0.99	7	6.5	1.08
4	58	52.3	1.11	2	7.7	0.26
5	74	68.8	1.08	5	10.2	0.49
6	64	60.8	1.05	6	9.2	0.65
7	67	71.5	0.94	16	11.5	1.39
8	54	63.6	0.85	22	12.3	1.79
9	56	59.5	0.94	20	16.5	1.21
10	37	43.2	0.86	101	94.8	1.06
Total	**571**	**571**	**1**	**184**	**183.9**	**1.00**

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	32.632	8	.000

The calibrative capacity of the Paediatric Index of Mortality 2 (PIM2) score was assessed by the Hosmer lemeshow goodness of fit test in which predicted mortality risk as assessed by PIM2 score was divided into ten deciles. It was compared with observed mortality. Table 14 shows the calibration of PIM2 score model of our study.Figure-12 shows the graphic representation of the calibration. The total observed mortality was 184 and the predicted mortality was 183.9 in our study with a standardized mortality ratio of 1.070 with 95% confidence interval 1.060 – 1.093. How ever closer look at the values show that the PIM2 score predicts more death in the low risk group compared to the observed mortality and under predicts mortality in the high risk group. The chi square value for our study was 32.6 with a p value of <0.001 suggesting poor calibration of PIM2 score in our study population.

Discussion

Intensive care treatment for critically ill children is being increasingly recognized as a need in India. It is no surprise therefore, that Paediatric Intensive Care is now a rapidly developing specialty with intensive care units being established in several places.

While those involved in the delivery of intensive care to critically ill children constantly strive to improve the standards of care given to a sick child, they are challenged by limited resources in terms of space, technology, manpower and finances. Admission into the intensive care unit often, is only the beginning of various other ICU interventions which can be difficult to limit once they are initiated. Such interventions come at a huge cost to the family and the outcome is uncertain despite all these interventions. Since the cost of health care in ICU has gone up dramatically, correctly identifying those children who are salvageable from moribund group of children becomes necessary for implementation and rationing of effective medical therapy.

This problem of scarce resources and the constant urge to utilize the resources optimally has led us to this study, to assess a scoring system which may guide us to prognosticate our patients. Additional benefit would be that this may enable us to compare the performance of our PICU with others.

Literature review revealed that there are two important scores available for paediatric use. Out of the two scores, the Paediatric Index Mortality 2 Score (PIM-2) is easy to collect since it has minimum number of variables; the data is to be collected at the time of admission to the intensive care unit or within the first hour; the software for calculating the PIM expected mortality is freely available on the internet for calculation. Though treatment provided in the emergency department could influence the outcome, the treatment provided in the PICU will not be influencing PIM 2 scoring system. To our knowledge, validation of Paediatric Index of Mortality 2 score has been done in only one study in Indian children but has been validated extensively elsewhere[40-44].

The study was done in PICU of Christian Medical College and Hospital, Vellore which is a major tertiary care centre in South India. The PICU has 11 beds and receives sick children with both medical and surgical diseases. The unit is busy with an admission rate of over 1200 children per year. The unit is equipped with ventilators, monitors and facilities for dialysis. Since this is a referral hospital's intensive care unit the patient population represents a good amount of case mix and disease severity requiring intensive care. The overall mortality lies around 20-25 % which is high compared to western figures. The reason for the high mortality could be because the children are sicker or because the quality of intensive care is not sufficient for the degree of illness. The patient population could also be very different in terms of the conditions that are prevalent here. The health seeking behaviour of the families when a child falls sick which depends on the cultural background from where they come from, the educational status, the finances, the primary health care accessible, the referral system available etc. could also influence the outcome. Hence we postulated that our children would be different from those in Australian ICUs and a scoring system developed in them may not perform well in our scenario.

During the study period, there were 755 children, aged from 72 hours to 15 years admitted to the PICU who were all included for the final analysis. The median age was 12 months with an interquartile range of 4-72 months. Infants made up 40 %(305) of the study population with neonates comprising 30% of them. This compares with the study population in the original PIM2 study done in Australian PICUs which had 42% infants. But in that study, the proportion of neonates was only 13%. The numbers of children in the rest of the age groups were also similar to original PIM2 study -33% compared to 28% in the 1-5year age group, 14% in both the 6-10yr and 11-15 yr age group as compared to 14 and 15% respectively in the Pearson et al study[28]. The age distribution of our study population was also similar to the only other Indian study done on PIM2 scoring system by Thukral et al[45]. The age distribution did not follow the normal distribution curve and this has been noted in other studies as well.

There were 435 (57.6%) boys and 320 (42.4%) girls with the male: female ratio of 1.4:1 showing preponderance for male gender. It was similar to the gender distribution of our hospital admissions. It may due to the cultural importance given to male gender for inheritance and seeking medical care for them more than for the girls. The study by Choi et al[41] from Hong Kong also showed a similar phenomenon.

Distribution of the study children according to the admitting units as described under results showed good amount of case mix in this study. As compared to the Pearson et al[28] study in which the proportion of children with surgical problems was 40%, our study showed that a significant number of children belonged to the medical units (92.5%). In their study 42% of the admissions were received from operating theater (OT), while only 2.8% of our population was received from the operating theatre. In our hospital cardiothoracic and neurosurgical units have their own separate ICUs where their post-operative patients are looked after and therefore they were not included in our study. This situation is very different from the ICUs included in the Pearson et al[28] study but not different from that described in the other Indian study by Thukral et al[45] and Turkey study by Ozer et al[40]. In general, PICUs in developing countries received children with medical problems mainly tropical infectious diseases seen in that part in significant numbers than in developed countries.

Sepsis with multiorgan dysfunction (18.8%) was the predominant reason for admission to our ICU. There could have been 2 reasons for this phenomenon. As we have seen earlier, 30% of admissions were in the neonatal age group and the most common diagnosis among them was sepsis. Apart from this, a dengue epidemic and an increased occurrence of Rickettsial infections that was encountered during the time the study was undertaken, also have contributed to significant number of admissions with shock and with MODS. Central nervous system illness (17.4%) was the second most common diagnosis in our study. It comprised of status epilepticus, viral encephalitis, Guillain Barre syndromes and severe bacterial meningitis. There was a slight reduction in the respiratory illnesses (12.8%). Acute gastroenteritis with shock and/or other complications contributed much to GI and hepatic illnesses (12.5%) apart from hepatic encephalopathy

and massive GI bleeds. Surprisingly poisoning and envenomation contributed 9.2%. Snake bite, scorpion sting envenomations, Organo-phosphorus poisoning, accidental ingestion of antipsychotic drugs, anti – histamines, camphor contributed significantly. Surgical emergencies and post operative monitoring for major surgical illnesses made up 4.1%. Acute lymphoblastic leukemia significantly contributed to malignancy (6.9%).

The distribution pattern of the diseases in our PICU was quiet different from the study by Pearson et al[28], where sepsis and other infectious diseases were not represented. The original PIM2 study group comprised mainly of surgical patients with low risk diagnoses according to PIM2 study authors themselves. However, a predominant sepsis and other infectious diseases pattern seen in our study is similar to others done in developing countries like Thukral et al in India (37.2%) and Ozer et al[40] in Turkey(41%). This represents the pattern of prevalence of diseases in developing countries which is quite different from developed countries. 12.8% of our children only had respiratory illness as the main diagnosis as compared to 21.6% reported by Pearson et al[22] (Australia), 14.97% by Pearson et al[28] (UK), 33% by Ozer et al (Turkey), 25.6 % by Thukral et al[45] (India), 39.6% by Choi et al[40] (Hong Kong). The likely reason for the low respiratory illnesses seen in our study population could be due to the fact that the study was completed just before the peak RSV season and/or due to the a reduction in the occurrence of severe ARIs related to the increasing utilization of Haemophilus influenzae b vaccine.

In our study, 254(33.6%) children out of 755 were ventilated within one hour of admission into PICU.

Out of 755 children, 571(75.7%) children survived and were discharged well from the intensive care unit. Of 184(24.3%) children who succumbed 104(13.8%) children died in the ICU and 80(10.5%) were discharged in a terminally ill condition at parental request. We have considered those who were discharged in a terminal condition as mortality. Out of the 184 children, 62(33.6%) had sepsis with multiorgan dysfunction group followed by neurological illness in 25(13.6%). Gastrointestinal- hepatic illness and

malignancy constituted 19(10.3%) each and respiratory illness in 16(8.7%). Only one child (0.5%) from the surgical unit with extensive burns had died. In the original PIM study by Pearson majority of the deaths was due to cardiac condition (24.4%), followed by injury (15.6%), respiratory illness (14.5%) and neurological illness (9.05%). Non cardiac post operative patients contributed to the 4.34% of mortality.

The overall mortality seen in our ICU is comparable to the figures reported by Thukral et al[45] from India (35.3%) and Ozer et al[40] from Turkey (27.6%). Other similar studies on prognostic scores however report much lower death rates: Pearson et al[28] (UK) 7.955%, Pearson et al[22] (Australia) 5.31%, Reinoul et al[42] (Netherlands) 6.6%, Martha et al[43] (Brazil) 7.83%, Choi et al[41] (Hong Kong) 2.64%.

Given the characteristics of our patients we did not expect the PIM to perform well in our intensive care unit and our null hypothesis was that PIM2 scoring system will not predict mortality in our Indian study population. However the analysis revealed interesting results.

Performance of PIM2 Score

The overall expected deaths as predicted by PIM 2 score was 183.9 with an area under the receiver operating characteristic curve (ROC) of 0.839 with a 95% CI of 0.804 and 0.874 which was statistically significant. This was comparable to the PIM2 score discrimination capacity achieved in Australian PICUs by Pearson et al with AUC of 0.90 (95% CI of 0.89 -0.91). Similar outcomes have been observed in other studies that have evaluated both PIM and PIM2. In the study by Thukral et al from Delhi, the AUC (95% CI) for PIM was 0.82(0.76-0.88) and for PIM2 0.81 (0.75-0.87). In a similar study on validation of PIM2 score by Eulmesekian et al[44] in Argentina on 1574 intensive care patients with 41 deaths the AUC was reported to be 0.89 (95% CI of 0.89 – 0.92) . Furthermore, recently wolfler et al[46] from Italy in their study validating PIM2 in 3266 children from 18 PICUs have reported an AUC of 0.89 (95% CI of 0.86 – 0.91). All these studies have concluded that both PIM and PIM 2 had adequate discriminative ability. The

only exception is the study by Ozer et al[40] in Turkey in which the AUC was only 0.69 (95%CI of 0.56 to 0.82) suggesting PIM had poor discriminatory capacity in their study population. However, the study was done in an 'Infantile intensive care unit' among children aged 1 month- 24 months, with the mortality of over 50%.

The standardized mortality ratio (SMR) calculated by PIM 2 score for our study group was 1.077 with 95% CI of 1.060 – 1.093. This was comparable to the SMR predicted in Australian PICU by Pearson et al[22] with a value of 1.00 (95% CI of 0.95 -1.05). It suggests that our PICU performance was comparable to the PICUs where the original PIM2 score was developed.

Similar outcomes have been observed in other studies that have evaluated both PIM and PIM2. The study on validation of PIM2 score by Eulmesekian et al in Argentina reported a SMR of 0.85 (95%CI 0.60 – 1.1). The earlier PIM validation in UK, Hong Kong, Netherlands reported SMRs of 0.87(95%CI 0.81-0.94), 0.61(95%CI 0.50-0.77), 0.88(95%CI 0.55-1.20) respectively. However SMRs reported from Turkey, Brazil and India have shown values of 3.68 (95%CI 3.08-4.28), 1.26(95%CI 0.87-1.77), 1.57(95%CI 1.24-1.59) respectively.

The calibration of PIM2 score as tested by Hosmer Lemeshow goodness of fit test for our study population showed a chi 2 value of 32.6 with a p value of <0.001, suggesting highly significant result. This means that the calibration of the PIM2 score was poor in our study population. Further analysis of the distribution of PIM2 score showed that the mortality was over predicted in the low risk categories and under predicted in the high risk categories. This may be attributable to the observation that only 97 of 755 (12.9%) children in the study population were allotted to risk groups either high or low risk. The high risk groups in the our scenario like sepsis, dengue hemorrhagic fever with dengue shock syndrome, Rickettsial infection in septic shock, cerebral malaria, scorpion sting, OP poisoning, rheumatic heart disease with multiple valvular lesions, hemolytic uremic syndrome, immuno suppression for chronic illness etc are not given enough high risk weightage in the PIM2 score. This could have led to the poor calibration

of PIM2 scoring with effect of under prediction of mortality in the high risk group. Similar effect of over prediction in the low risk group could have produced an opposite effect.

Poor calibration of PIM and PIM2 has been reported in other studies done earlier elsewhere. Although the study done by Choi et al from Hong Kong reported satisfactory overall calibration of PIM , there was an over estimation of mortality in 0% - <25% and 50 -<75% risk category. Conclusion could not be drawn in the other risk categories in view of wide range of 95% CI for SMR and insufficient sample size. Ozer et al from Turkey reported poor discrimination of PIM in their infantile intensive care unit, but satisfactory overall calibration of PIM with a chi 2 value of 10.9 and p value of 0.20. Reinoud et al from Netherlands reported an underestimation of mortality among the low risk group (1 -<5%), overestimation in >30% risk category with an overall satisfactory calibration (chi 2 4.92 p=0.77) in their study population. Martha et al from Brazil and Eulmesekian et al from Argentina have also reported poor calibration of PIM2 score among their study population. The only Indian study done by Thukral et al in New Delhi revealed a chi 2 value of 7.64 (p=0.47), however 95% CI of SMRs were wide. How ever studies from UK (PIM), Italy on PIM 2 had good calibration. This could be due to major surgical population in the developed countries in contrast to severe medical conditions seen in the developing countries. The higher SMR in developing countries may be attributed to prevalence of infectious diseases, predominantly sepsis as major cause of admission, malnutrition, chronic illness, delayed referral, absence of good transport facilities for sicker children, scarce resources, both physical and human staffing of PICUs and differences in the practice guidelines in the various PICUs.

Limitations of the study

In our study base excess analysis was not done to all unless clinically warranted. Arterial blood gas analysis was routinely done only for those patients who were intubated on arrival and subsequently ventilated. Some of the blood gases were done after the first hour and were not included in the study. As we did not have an Fiao2 analyzer, those who were on high flow oxygen could not be assessed for value of Fio2 and the variable in

PIM 2 scoring system was entered as 0. This could have underestimated the severity of the illness. During our study period, there was a Dengue epidemic and increased prevalence of Rickettsial infection. Though these patients appeared stable initially, some of them later went into refractory shock with ongoing third space loss fluid loss and haemorrhage. This group of children could have had an under estimation of their predicted mortality.

Our study shows that PIM had good discriminatory capacity in our patient population and therefore given its advantages is a good scoring model applicable for our patients. However, we observed that the PIM2 did not calibrate well with an under

prediction of mortality in the high risk group and over prediction in the low risk group. Further studies are needed to identify high risk diagnosis in our population. This can be followed up with studies that evaluate PIM2 with those diagnoses coded as high risk in the PIM2 model. Other areas of research would be to determine if PaO2 could be substituted with Sa02 in the PIM2 variable 100*FiO2/PaO2, this would further improve entries of this particular PIM2 variable. The possibility and predictability of PIM2 score with these modified variables should be explored in the future studies. This may result in modifying or adapting the existing PIM 2 score in a way that their current functioning in the developed world is not affected but are appropriate for use within the developing world.

Conclusions

- PIM 2 score performed well in predicting deaths well among the children admitted to the Paediatric Intensive Care Unit.

- The Standardized Mortality Ratio of the study population was 1.077 (95% CI 1.060 – 1.093) showing the performance of the PICU to be comparable to the Australian PICUs where PIM2 was developed.

- PIM2 model calibrated poorly in our study population as tested by Hosmer Lemeshow goodness of fit test.

- Further studies evaluating a modified and adapted PIM2 model that incorporates the different high risk diagnoses seen in our population are needed for better calibration.

Bibliography

1. Martin HK, Schuster DP. Predicting intensive care unit outcomes with scoring systems. Underlying concepts and principles. Crit Care Clin 1994 10(1):1-14.
2. Miller DH. The rationing of intensive care. Crit Care Clin 1994 10(1):135-144.
3. Keene AR, Cullen DJ. Therapeutic intervention scoring systems: Update Crit Care Med 1983 11(1) 1-3.
4. Knaus WA, Draper EA, Wagner DP, et al. APACHE II: A severity of disease classification system. Crit Care Med 1985 13(8):18-29.
5. Le Gall JR , Loirat P, Alpertovich A et al. A Simplified acute physiology scores for ICU patients. Crit Care Med 1984 1(2):975 to 977.
6. Gunning K, Rowan K. ABC of intensive care. Outcome data and scoring systems. Clinical review BMJ 1999 24 Jul (319):241-244.
7. Henderson A J **,** Garland L **,** Warne S**,** Bailey L **,** Weir P**,** Edees S**.** Risk adjusted mortality of critical illness in a defined geographical region. Arch Dis child 2002 86:194-199.
8. Copes WS, Sacco WJ, Champion HR, Bain LW, "Progress in Characterising Anatomic Injury", In Proceedings of the 33rd Annual Meeting of the Association for the Advancement of Automotive Medicine, Baltimore, MA, USA 205-218.
9. Lavoie A, Moore L, LeSage N, Liberman M and Sampalis J. The Injury Severity Score or the New Injury Severity Score for predicting intensive care unit Injury. 2005 36(4):477-483.
10. Gray JE, Richardson DK et al. Neonatal Therapeutic Intervention Scoring System: a therapy-based severity-of-illness index. Pediatrics. 1992 (90):561-7.
11. The International Neonatal Network, The CRIB (clinical risk index for babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. Lancet 1993 (342**)** pp. 193–198.
12. Leteurtre S, Martinot A, Duhamel A, et al. Validation of the pediatric logistic organ dysfunction (PELOD) score. A prospective multi-center study. Lancet 2003 362:192–197.

13. Richardson DK et al. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. Pediatrics. 1993 91:617-23.
14. Miranda DR et al. Simplified Therapeutic Intervention Scoring System: the TISS-28 items. Results from a multi-center study. Crit Care Med. 1996 24:64-73.
15. Cowen JS, Kelley MA. Errors and bias in using predictive scoring systems. Crit Care Clin 1994 10(1): 53-72.
16. Zimmerman J E, Kramer A A, Mcnair d S, Malila F M. Acute. Physiology and Chronic Health Evaluation (APACHE). IV: Hospital mortality assessment for today's critically ill patients. Crit. care med. 2006 34(5):1297-1310.
17. Pollack MM .Clinical scoring systems in pediatric intensive care. In: Fuhrman BP, Zimmerman JJ (eds) Pediatric critical care. Mosby Year Book, St. Louis, pp 153–162.
18. Lacroix J, Cotting J. Severity of illness and organ dysfunction scoring in children; for the Pediatric Acute Lung Injury and Sepsis. Pediatr Crit Care Med 2005 Vol. 6, No. 3 Suppl.
19. Pollack MM, Ruttimann UE, Getson PR. Pediatric risk of mortality (PRISM) score 1988. Crit Care Med. 16(11):1110-1116.
20. Pollack G, Murray M, Kantilal M, Ruttiman, Urs E. PRISM III: An updated Pediatric Risk of Mortality score. Crit Care Med 1996 (24):743-752.
21. Yeh TS, Pollack MM, Ruttimann UE, Holbrook PR, Fields AI Validation of a physiologic stability index for use in critically ill infants and children. Pediatr Res 1984 18:445–451
22. Shann F, Pearson G, Slater A, Wilkinson K, Paediatric index of mortality (PIM): a mortality prediction model for children in intensive care, Intensive Care Med 1997 23:201–207.
23. Randolph A.G. Paediatric index of mortality (PIM): do we need another paediatric mortality prediction score? Intensive Care Med 1997 23:141–142.
24. Copas JB. Plotting p against x. Applied statistics 1983 32:25–31.
25. Hosmer DW, Lemeshow S. The effectiveness of the transformation checked with other statistical methods. Applied logistic regression 1989 Wiley, New York.

26. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules: applications and methodological standards. N Engl J Med 1985 313:793–799.
27. Slater A, Shann F, Gale P for the PIM Study Group PIM2: a revised version of the Paediatric Index of Mortality. Intensive Care Med (2003) 29:278–285.
28. Pearson GA, Stickley J, Shann F. Calibration of the paediatric index of mortality in UK paediatric intensive care units. Arch Dis Child 2001 84:125-128.
29. Graciano AL et al. The Pediatric Multiple Organ Dysfunction Score (P-MODS) : development and validation of an objective scale to measure the severity of multiple organ dysfunction in critically ill children. Crit care Med. 2005 Jul; 33(7):1484-91.
30. Hanley JA, McNeil BJ.A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983 148:839–843.
31. Ruttimann UE, Pollack MM. Objective assessment of changing mortality risks in pediatric intensive care patient. Crit Care Med 1991 19:474-83.
32. G. Parry et al. CRIB II: an update of the Clinical Risk Index for Babies scores. Lancet 2003 361(9371):1789-91.
33. Richardson DK et al. Score for Neonatal Acute Physiology: a physiologic severity index for neonatal intensive care. Pediatrics. 1993 91:617-23.
34. D K. Richardson et al. SNAP-II and SNAPPE-II: Simplified newborn illness severity and mortality risk scores. J Pediatr 2001 138: 92-100.
35. TEPAS JJ et coll. The Pediatric Trauma Score as a predictor of injury severity in the injured child. J. Pediat. Surg. 1987 22:14-8.
36. Rowan KM, Kerr JH, Major F, McPherson K, Short A, Vessey MP Intensive care society's APACHE II study in Britain and Ireland. In: Variations in case mix of adult admissions to general intensive care units and impact on outcome. BMJ 1993 307:972–977.
37. Dragsted L, Jorgensen J, Jensen N, Bonsing E, Jacobsen E, Knaus WA, Qvist J. Inter hospital comparisons of patient outcome from intensive care: importance of lead-time bias. Crit Care Med 1989 17:418–422.

38. Castellanos-Ortega A et al. A new prognostic scoring system for meningococcal septic shock in children. Comparison with three other scores. Intensive Care Med 2002; 28:341-351.
39. Sinclair JF et al. Prognosis of meningococcal septicaemia. *Lancet* 1987; 2(8549):38.
40. Ozer et al. The Comparison of PRISM and PIM Scoring Systems for Mortality Risk in Infantile Intensive Care. Journal of Tropical Pediatrics, 2004; Vol. 50, No. 6, 334-338.
41. Choi et al. Assessment of the Paediatric Index of Mortality (PIM) and the Paediatric Risk of Mortality (PRISM) III score for prediction of mortality in a paediatric intensive care unit in Hong Kong. Hong Kong Med J 2005; Vol 11: No 97- 103.
42. Reinoud.J.B.J.Gemke et al. Scoring systems in pediatric intensive care: PRISM III versus PIM. Intensive Care Med 2002 28:204–207.
43. Martha et al. Comparison of two prognostic scores (PRISM and PIM) at a pediatric intensive care unit. Jornal de Pediatria 2005; Vol. 81, No.3, 259-264.
44. Eulmesekian PG, Pérez A, Minces PG, Ferrero H. Validation of pediatric index of mortality 2 (PIM2) in a single pediatric intensive care unit of Argentina: Paediatr Crit Care Med. 2007 Jan;8(1):54-7.
45. Thukral et al. Performance of Paediatric Risk of Mortality (PRISM) and Paediatric Index of Mortality (PIM), and PIM2 in a pediatric intensive care unit in a developing country. Paediatr Crit Care Med 2006; Vol 7, No. 4, 356 – 361.
46. Wolfler, A; Silvani, P; Musicco, Massimo; Salvo, I,: Pediatric Index of Mortality 2 score in Italy: a multicenter, prospective, observational study , Intensive Care Medicine 2007, Aug, 33, 8, 1407-1413.

Appendices

APPENDIX: 1

Paediatric index of mortality 2 – instructions for use[27]

PIM2 is calculated from the information collected at the time a child is admitted to your ICU. Because PIM2 describes how ill the child was at the time you started intensive care, the observations to be recorded are those made at or about the time of first face-to-face (not telephone) contact between the patient and a doctor from your intensive care unit (or a doctor from a specialist paediatric transport team). Use the first value of each variable measured within the period from the time of first contact to 1 h after arrival in your ICU. The first contact may be in your ICU, your emergency department, a ward in your own hospital, or in another hospital (e.g. on a retrieval). If information is missing (e.g. base excess is not measured) record zero, except for systolic blood pressure, which should be recorded as 120. Include all children admitted to your ICU (consecutive admissions).

Coding rules. These rules must be followed carefully for PIM2 to perform reliably:

1. Record SBP as 0 if the patient is in cardiac arrest, record 30 if the patient is shocked and the blood pressure is so low that it cannot be measured.
2. Pupillary reactions to bright light are used as an index of brain function. Do not record an abnormal finding if this is due to drugs, toxins or local eye injury.
3. Mechanical ventilation includes mask or nasal CPAP or BiPAP or negative pressure ventilation.
4. Elective admission. Include admission after elective surgery or admission for an elective procedure (e.g. insertion of a central line), or elective monitoring, or review of home ventilation. An ICU admission or an operation is considered elective if it could be postponed for more than 6 h without adverse effect.
5. Recovery from surgery or procedure includes a radiology procedure or cardiac catheter. Do not include patients admitted from the operating theatre where recovery from surgery is not the main reason for ICU admission (e.g. a patient with a head injury who is admitted from theatre after insertion of an ICP monitor; in this patient the main reason for ICU admission is the head injury).
6. Cardiac bypass. These patients must also be coded as recovery from surgery.
7. Cardiac arrest preceding ICU admission includes both in-hospital and out-of-hospital arrests. Requires either documented absent pulse or the requirement for external cardiac compression. Do not include past history of cardiac arrest.
8. Cerebral haemorrhage must be spontaneous (e.g. from aneurysm or AV malformation). Do not include traumatic cerebral haemorrhage or intracranial haemorrhage that is not intracerebral (e.g. subdural haemorrhage).
9. Hypoplastic left heart syndrome. Any age, but include only cases where a Norwood procedure or equivalent is or was required in the neonatal period to sustain life.

10. Liver failure - acute or chronic must be the main reason for ICU admission. Include patients admitted for recovery following liver transplantation for acute or chronic liver failure.
11. Neuro-degenerative disorder. Requires a history of progressive loss of milestones or a diagnosis where this will inevitably occur.
12. Bronchiolitis. Include children who present either with respiratory distress or central apnoea where the clinical diagnosis is Bronchiolitis.
13. Obstructive sleep apnoea. Include patients admitted following adenoidectomy and/or tonsillectomy in whom obstructive sleep apnoea is the main reason for ICU admission (and code as recovery from surgery).

Appendix 2: PIM2 Study Proforma

A STUDY ON THE VALIDATION OF PAEDIATRIC INDEX OF MORTALITY 2 (PIM2) SCORE IN THE PEDIATRIC INTENSIVE CARE UNIT OF CMC, VELLORE

Name : Age: Hospital No:-

Date of admission: . . Time:

Admission from:- Ward / Casualty / Operation Theatre / Other

Unit:- CH I / CH II / CH III / Paed Surgery / Haematology / Paediatric Oncology / Paed Nephro / Orthopaed / Trauma / Plastic surgery / ENT / Others

Diagnosis :-

Length of stay:-

PIM 2 SCORE

Paediatric index of mortality

Variables	Values (1 – yes, 0 – no)
Elective admission	
Recovery post procedure	
Cardiac bypass	
High risk diagnosis	
Low risk diagnosis	
No pupillary response to bright light (>3 mm and both fixed)	
Mechanical ventilation (at any time during the first one hour)	
Systolic blood pressure (mm of Hg)	
Base Excess (mm of Hg) (arterial or capillary blood)	
FiO2*100/PaO2(mm of Hg)	

Outcome :

Death
Discharge at request
Transferred to Ward

Appendix 3: Online PIM 2 calculator

http://www.sfar.org/scores2/pim22.html

PIM 2 **(Paediatric Index of Mortality)**

Variables (help)	**Values** (1 if Yes, 0 otherwise)	**Beta**
Elective admission		
Recovery post procedure		
Cardiac bypass		
High risk diagnosis		
Low risk diagnosis		
No response of pupils to bright light (> 3 mm and both fixed)		
Mechanical ventilation (at any time during first hour in ICU)		
Systolic Blood Pressure (mmHg)		0.01395
Base Excess (mmHg) (arterial or capillary blood)		0.1040
FiO2*100/ PaO2 (mmHg)		0.2888

Predicted Death Rate :

Clear

Logit = (-4.8841) + (values * Beta) + (0.01395 * (absolute(SBP-120))) + (0.1040 * (absolute base excess)) + (0.2888 * (100*FiO2/PaO2))

Predicted death rate = $e^{Logit}/ (1+e^{Logit})$

Appendix 4: Coding for the master Sheet

A: Name of the patient (Name)
B: Sex of the patient (sex)
C: Age of the patient in years (age)
D: Age of the patient in months (Months)
E: Time of Admission (TOA)
F: Hospital Number (H.No)
G: Date of admission (DOA)
H: Date of Discharge (DOD)
I: Length of stay in PICU (Length of stay)
J: Source of admission (FROM)
K: Admitting Unit (UNIT)
L: Elective admission(EA) (1-yes,0-no)
M: Recovery post procedure (RPP) (1-yes, 0-no)
N: Post cardiac bypass (CB) (1-yes,0-no)
O: High risk diagnosis at admission (HIGH RISK) (1-yes,0-no)
P: Coding for the High risk (risk No)
Q: Low risk diagnosis at admission (LRD) (1-yes,0-no)
R: Coding for the low risk (risk No)
S: No pupillary response (No response of pupils to bright light (> 3 mm and both fixed)) (1-yes,0-no)
T: Mechanical ventilation (at any time during first hour in ICU) (MV) (1-yes,0-no)
U: Systolic Blood Pressure (mmHg) (Sys.BP)
V: Base Excess (mmHg) (arterial or capillary blood) (BASE EXCESS)
W: FiO2*100/ PaO2 (mmHg) (OI)
X: Predicted Death Rate (PIM2)
Y: Outcome of the patient (OUTCOME) (DAMA,TO,DEATH)
Z: Coding for the Outcome (outcome no) (1- Death/DAMA,0-TO)
AA: Primary Diagnosis (PRIMARY DIAGNOSIS)
AB: Coding for the primary diagnosis (coding)
AC: Secondary diagnosis (SECONDARY DIA)
AD: Coding for the secondary diagnosis (CODING)
AE: Tertiary diagnosis(TERTIARY DIAGNOSIS)
AF: Primary system of illness(PRIMARY SYSTEM OF ILLNESS)
AG: Categorization of the PIM2 into five groups(1 to 5)